2003

An incomplete contract analysis of multinational enterprises and oil field unitization

Vijay Mohan
Iowa State University

An incomplete contract analysis of multinational enterprises and oil field unitization

by

Vijay Mohan

A dissertation submitted to the graduate faculty

in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Major: Economics

Program of Study Committee:
Harvey Lapan, Major Professor
Samuel DeMarie
Brent Hueth
John Schroeter
Jinhua Zhao

Iowa State University

Ames, Iowa

2003

UMI Number: 3105093

Graduate College
Iowa State University

This is to certify that the doctoral dissertation of

Vijay Mohan

has met the dissertation requirements of Iowa State University

Signature was redacted for privacy.

Major Professor

Signature was redacted for privacy.

For the Major Program

To my parents

For always believing

TABLE OF CONTENTS

I. INTRODUCTION: INCOMPLETE CONTRACTS AND THE PROPERTY RIGHTS THEORY

1. Theories of the firm

The growing discontent among economists with the ability of the neoclassical theory to handle various issues related to the behavior of a firm has resulted in new theories being developed over the past few decades that attempt to fill this void. The neoclassical theory relies heavily on the technological aspects of production for its foundations, and a firm is, in essence, treated as a "black box". Inputs are converted to outputs by the firm on the basis of some exogenously given engineering technology, and this process is carried out by an owner (or manager) whose sole responsibility is to take decisions that ensure the maximization of profits for the firm.

In spite of (or perhaps because of) the neoclassical theory's disregard for a precise formulation of the underlying structure of a firm, it has served a very useful purpose in broadening our understanding of how the economy works in general, and the behavior of industries, the interaction between the producers and consumers, the analysis of varying degrees of market power, and the evaluation of public policy, in particular.

The treatment of the firm in the neoclassical theory does unarguably, however, have its limitations and there are two prominent criticisms that have been leveled against it[1]. The first relates to the fact that the neoclassical theory has very little to say about the internal organization of the firm. A firm is often a complex hierarchical structure; owners delegate responsibility to a board of directors, who hire managers to oversee the activities of the firm, and the managers, in turn, allocate tasks to employees. An understanding of the nature of the interactions between the various tiers of this hierarchy is essential to accurately model the behavior of the firm, and the endeavor to address this issue has resulted in the development of the principal-agent theory.

There are two underlying ideas that have been used to construct this theory; the first is the existence of a conflict of interest between two individuals (the principal and the agent), and the second is the presence of informational asymmetries. This asymmetry in information could take the form of an action that the agent takes that is unobservable to the principal

[1] See Hart (1995, Chapter 1) for a more elaborate discussion.

(moral hazard), or of the agent possessing some information that the principal does not have access to (hidden information or adverse selection). In either case the agent can take advantage of the asymmetry in information to further her own interests, to the detriment of the principal. In the presence of informational asymmetries, the principal will devise a contract that provides incentives for the agent to alter her behavior in a manner that benefits the principal.

Building on this intuitive line of reasoning, principal-agent models have been able to provide remarkable insights into the incentive structures prevalent in organizations[2]. It is worth noting that the contracts themselves are complete in the framework of the principal-agent theory: any variable that is observable can be contracted upon costlessly, and so contracts will incorporate comprehensive descriptions of the actions of the contracting parties for any future contingency.

The second criticism that has been put forward against the neoclassical theory (which the principal- agent theory does not resolve satisfactorily) is the failure of the neo-classical firm to pinpoint the boundaries of a firm. The inadequacy of the neoclassical theory in providing a precise definition of what constitutes a firm was first pointed out by Coase (1937). Coase's arguments revolve around his observation that market transactions are coordinated through the price mechanism, whereas within a firm the price mechanism is replaced with a power structure where an entrepreneur coordinates the actions of the employees. A firm can carry out a transaction either through the market, or it can integrate the transaction within the firm. Any transaction in the market which employs the price mechanism involves certain costs, such as costs associated with discovering what the relevant prices are, the costs of negotiating and concluding a contract for each market transaction, and costs arising out of the inability of agents to forecast the future accurately, which poses difficulties for agents when entering into long term contracts[3]. On the reverse side, incorporating an additional transaction into the firm has costs associated with it as well,

---

[2] See Hart and Holmstrom (1987) for a survey, and Salanie (1997) and Macho-Stadler and Perez-Castrillo (1997) for an extensive treatment.

[3] The idea behind the last case is that the inability of agents to forecast the future results in contracts being incomplete, and the longer the duration for which the contracts lasts, the harder it is for agents to describe commitments of the parties for the future. As the relationship unfolds, one agent then has more power in directing the other over the use of resources (the hold up problem). This example is interesting because it shows Coase's influence on the theories that were to follow.

which consist mainly of decreasing returns to the activities of the entrepreneur. Coase then postulated that a firm would expand by organizing additional transactions within the firm till the costs are equalized at the margin.

Coase's pioneering efforts paved the way for the formulation of the transaction cost theory of the firm. Williamson (1979, 1985) and Klein, Crawford and Alchian (1978) are widely acknowledged for transforming Coase's intuition into a concrete theory. The transaction cost literature focuses on the costs of writing and enforcing contracts, which lead to contracts being incomplete. Transaction costs exist due to three reasons[4]. First, agents have a limited ability to forecast the future and anticipate all contingencies that can occur: they have bounded rationality. Second, even if agents are perfectly rational, there may be costs involved in negotiating an acceptable contract. Third, there are costs associated with writing a verifiable contract that can be examined by a court in case of a dispute. The incompleteness of contracts necessitates renegotiation of the contract as states of the world are revealed to agents. Williamson (1979) identifies three dimensions that characterize transactions: the level of uncertainty, the frequency with which a transaction recurs, and the degree to which assets used for the transaction are relationship specific[5]. The final link between contractual incompleteness, the characteristics of a transaction and the institutional mode that governs the transaction is the assumption that agents are opportunistic. The repercussions of opportunistic behavior are especially evident when there is a high degree of relationship specificity, and this gives rise to the classic "hold up" problem.

When contracts are incomplete, the parties' inability to specify their obligations for all future contingencies forces them to renegotiate the initial contract as and when relevant states of the world are realized. The presence of relationship specific assets, however, results in the development of *ex post* quasi-rents that arise from the fact that assets are more valuable inside the trade than outside (Klein, Crawford and Alchian (1978)). When it comes to

[4] There is no specific definition of a transaction cost in the literature, making the concept a slippery one to handle. Indeed, Fischer (1977) writes, "...there is a suspicion that almost anything can be rationalized by invoking suitably specified transaction costs."

[5] Williamson lists four types of relationship specific assets: site specific assets, dedicated assets, relationship specific physical assets and relationship specific human assets. The first and the last are self-explanatory. The difference between a dedicated asset and a relationship specific physical asset is that the former involves the agents building productive capacities for which there is insufficient demand outside the relationship, while the latter relates to a situation where parties make investments in specialized physical assets.

renegotiating the contract each agent, being opportunistic, will try to "hold up" the other to extract as much of the *ex post* quasi-rent as possible. Integration is then the solution adopted by one party to mitigate the hold up problem.

Williamson (1979) explains the alleviation of the hold-up problem within a firm by noting that "...[under vertical integration] adaptations can be made in a sequential way without the need to consult, complete, or revise inter-firm agreements." This statement reflects Coase's idea that the distinguishing feature of a firm is the existence of some figure of authority within the firm who directs the actions of the employees. An implicit assumption of this argument, of course, is that employees are willing, within reasonable limits, to follow the directions of the employer.

Alchian and Demsetz (1972) question this line of reasoning by arguing that the firm has no more authority when dealing with its employee than it does when it interacts with an independent agent in the market; what distinguishes a firm in their view is its ownership of a bundle of decision rights[6]. Alchian and Demsetz provide the following example to establish the first claim. If a consumer is unhappy with his grocer (or if a firm is unhappy with an independent agent), the worst she can do is to take her custom elsewhere, that is, the consumer can "fire" the grocer. But that is usually the option that the employer has in dealing with a disobedient employee. As Tirole (1999) points out, the fact that the firm has the common right to fire its employees and an independent agent does not necessarily imply that the two situations are similar, and is in itself, therefore, not damaging to the Coase-Williamson view that authority relationships exist within a firm. Their second claim that the owner of a firm possesses a bundle of rights[7] is, however, a very useful insight and forms a part of the foundations of the property rights theory. As Tirole states, "The two concepts of authority (decision right, control right) and property rights are closely related and therefore sometimes confused...a property right is a bundle of decision rights." This point is further elaborated upon in the next section.

Grossman and Hart (1986) proffer more damaging criticisms of the transaction-cost

[6] According to Alchian and Demsetz, these rights are: (1) to be a residual claimant; (2) to observe input behavior; (3) to be the central party common to all contracts with inputs; (4) to alter the membership of the team; and (5) to sell these rights.

[7] The same observation was made by Walras as early as the turn of the century.

theory. First, they observe that according to the transaction cost literature, opportunistic behavior arising during renegotiation of an incomplete contract is the main motivation for firms to integrate. The theory does not, however, explore how such opportunistic behavior changes when formerly independent agents become employees. Secondly, they present the following argument to establish the fact that the transaction-cost literature does not really succeed in demarcating the boundaries of the firm: "… if vertical integration always reduces transaction costs, any buyer A and seller B that have a contractual relationship should be able to make themselves better off as follows: (i) A buys B and makes the previous owner of B the manager of the new subsidiary; (ii) A sets a transfer price between the subsidiary and itself equal to the contract price that existed when the firms were separate enterprises; and (iii) A gives the manager of B a compensation package equal to the profit of the subsidiary. Given this, however, how can integration ever be strictly worse than non-integration; that is, what limits the size of the firm?"

Coase's retort to this would have been that there are costs to organizing another transaction within a firm, mainly the decreasing returns to entrepreneurial function (a manager has limited capacity to efficiently control too many activities), and that these costs increase as additional transactions are organized within the firm. This appears plausible, but as Evans and Grossman (1983) point out, the firm could always hire another manager.

The property rights theory has sprung up from the dissatisfaction with other theories in pinpointing the limits of a firm, and from a desire to develop an analytically precise paradigm (the transaction-cost literature relies very little on quantitative tools) that can provide insights into the factors that motivate firms to integrate and those that govern the internal organization of a firm.

2. The property rights theory

2.1 The methodology of the property rights approach

In an influential paper, Grossman and Hart (1986) set out the general framework for the property rights approach to analyzing the behavior of a firm, and in doing so opened the gates for a flood of research focused on the role of incomplete contracts in characterizing the behavior of firms. The property rights theory shares some of the features of the transaction cost literature, so it is worthwhile to outline the similarities between the two. First of all, both theories assume the prevalence of incomplete contracts, that is, the presence of transaction costs that impediment the comprehensive description of all contracting parties' future obligations. The reliance on the existence of *ex post* quasi rents (or gains from trade) due to relationship specificity is another common feature, and the opportunistic behavior of agents is crucial to the development of both theories.

The last item also serves as a point of departure between the transactions cost framework and the property rights theory. As mentioned before, the transaction cost literature does not specify how opportunism affects the firm. The property rights theory, on the other hand, allows for similar opportunistic behavior to occur in integrated firms and in market transactions. Moreover, the property rights theory focuses attention on how different organizational forms affect *ex ante* non-contractible relationship specific investments[8].

If contracts are incomplete, parties will be forced to renegotiate the terms of trade when the relevant state of the world becomes known. However, by then the parties are locked into a relationship and how the gains from maintaining that relationship are split depends on their bargaining powers during renegotiation; each party has an incentive to hold up the other one in order to extract as much of the *ex post* quasi rents as possible. Thus far the story is familiar: it follows along the lines of the transaction cost literature. The central feature of the property rights theory is that anticipating the occurrence of the hold up problem, each party will fear the appropriation of the returns on its *ex ante* investments by the other. This provides an incentive for parties to reduce the level of relationship specific investments, making investments instead that are less specific to its trading partner. Integration does not solve this problem in its entirety; *ex ante* investment levels are distorted in all organizational forms. If one firm acquires another, its incentives to increase investment would probably rise as it fears less appropriation of its investment by the other firm. The incentives of the acquired firm to invest are lowered, however, for exactly the opposite reason. The attention to *ex ante* decisions is the main feature that distinguishes the property rights theory from the transaction cost literature[9].

---

[8] The transaction cost literature recognizes the effect on ex ante investments as well, but not in such precise terms as the property rights theory.

[9] For an analysis of the implications of the difference in terms of empirical modeling see Whinston (2001).

We now address the issue of why property rights matter. Consider a physical, or more generally a non-human, asset owned by a firm[10]. The question that comes to mind is that in what sense does a firm "own" this asset, or put slightly differently, what exactly is the owner of an asset entitled to? In the previous section, the notion was introduced that a property right is a vector of control, or decision, rights. A contract between two parties allocates some of these control rights among the contracting parties, and the rights that are described in the contract are referred to as specific rights. An incomplete contract, however, has some missing gaps: there are states of the world where some decision right has not been allocated. The property rights approach takes the view that the owner of an asset is entitled to these *residual* control rights[11]: "...the right to decide all usages of the asset in any way not inconsistent with a prior contract, custom, or law" (Hart, 1995).

The formal modeling of the ideas behind the property rights approach is developed in Grossman and Hart (1986), Hart and Moore (1990) and Hart (1995). Though these papers vary in details, the underlying thread of logic is the same, and it is worthwhile to elaborate on the methodology that forms the basis of the property rights theory. For our exposition, we use the model presented in Hart (1995).

The standard model is usually framed in a three-date, two-firm and two-asset structure. At Date 0, a contract is signed that allocates property rights between the two (risk-neutral) firms; at Date 1, non-contractible relationship specific human capital investment decisions are made; at Date 2, exchange of the good and renegotiation of the original contract takes place. In the basic model, the contract is assumed to be a short term one: at Date 0, there is uncertainty about the exact nature of the good that will be required at date 2 (it depends on the relevant state of nature realized at Date 2), and if there are a large number of states of nature that can be realized at Date 2, it is prohibitively expensive to describe them in advance[12]. Long term contracts that prescribe agreements about the trade, such as the price of the good to be traded, are therefore not feasible. This forces parties to renegotiate the

[10] A non-human asset could be, for example, a patent, a brand name or simply a list of clients.
[11] Hart and Moore (1990) distinguish between ownership in terms of entitlement to an asset's verifiable profit stream and ownership in the sense of possession of residual control rights. They point out that the two may often, but not necessarily, go together. Tirole (1999) provides an example: an employee who has no control right may be entitled to a portion of the verifiable profit stream in the form of a bonus.
[12] Recall that that writing contracts involves transaction costs; even in the presence of transaction costs, the assumption that *no* aspect of trade is describable is an extreme one.

contract at Date 2, once the relevant state of nature has been realized[13]. There are two key assumptions that are made at this point. The first is that the firms have symmetric information regarding all variables. All investments, revenues and costs are assumed to be observable by the firms, but not verifiable by outside arbitrators, which renders them non-contractible[14]. The second is that parties are rational enough to be able to foresee correctly the payoffs associated with their actions.

The Date 0 contract, therefore, specifies only an allocation of the two assets between the two firms. Hart enumerates three 'leading' ownership structures that can be assigned at Date 0: non-integration (each firm owns one asset), type 1 integration (firm 1 owns both assets) and type 2 integration (firm 2 owns each asset). Given a rule for sharing the Date 2 (*ex post*) surplus, such as the Nash bargaining solution, rationality of the firms ensures that one can derive the Date 1 (*ex ante*) surplus that is associated with each property right allocation. The theory then predicts that the property right allocation that maximizes the *ex ante* total surplus is the one chosen by the two firms at Date 0. The symmetry in information between the two firms is prevalent both at Date 1 and at Date 2, and most existing models in incomplete contracts assume this. The assumption is largely one of convenience because it precludes the necessity for bargaining under asymmetric information.

In order to fix ideas, a brief outline of the model in Hart (1995) is provided here. There are two managers M1 and M2 who work with two assets a1 and a2. M2 uses the non-human asset a2 to produce an input, which following Hart, we shall call a widget. M1 requires a widget and the asset a1 to produce (one unit of) output for the final market. M1 can buy either a specific widget that is particularly suited to her needs from M2 or a non-specific widget from an outside supplier. Similarly, M2 can either produce a specific widget for M1 or supply a non-specific widget to the spot market. At Date 1, M1 and M2 make relationship specific investments, i and e, respectively, which are investments in human capital. If trade between the two occurs at Date 3, each manager has access to both assets and the human capital of the other manager. If trade does not occur the managers have to

[13] Maskin and Tirole (1999b) argue that the parties' inability to commit not to renegotiate is one of three crucial assumptions on which the foundations of incomplete contracts rests.
[14] This is in contrast to agency models where a variable is contractible if it is observable and an observable variable is assumed to be verifiable.

turn to the spot market, in which case each manager has access only to the asset that she owns.

The human capital investment of M1 affects the revenues generated by the sale of the output both within and outside her relationship with M2; similarly, M2's investment affects the cost of producing the input within and outside her relationship with M1. Keeping this in mind, we denote M1's revenues if trade occurs as $R(i)$ and her revenue if trade with M1 fails as $r(i;A)$, where the lower-case r represents the absence of M1's human capital, and A is the set of assets owned by M1 under a given property right allocation at Date 0. Likewise, $C(e)$ denotes M2's costs if she trades with M1, and $c(e;B)$ represents her costs when supplying the non-specific input to the spot market, where the lower-case c indicates the absence of M1's human capital, and B depicts the set of assets owned by M2. If trade occurs between M1 and M2, they negotiate a transfer p, and the total date 2 surplus under trade is given by $R(i)-C(e)$. If the (competitive) spot market is utilized in the absence of trade, where the price of the non-specific input is $\overline{p}$, the total date 2 surplus is $r(i;A)-c(e;B)$.

The fact that investments are relationship specific is modeled with the assumption that there always exist gains from trade, that is, $R(i)-C(e)>r(i;A)-c(e;B)\geq 0$. Given that the relationship is worth more than interacting with the spot market, M1 and M2 will always trade in equilibrium. A crucial feature of the model is the way the ex post gains from trade, $[R(i)-C(e)]-[r(i;A)-c(e;B)]$, are split up. Specifically, Hart (and many other papers in the literature) assumes that the gains from trade are split 50/50 through Nash bargaining. M1 and M2, therefore, will receive the following *ex post* payoffs:

(1) $$\Pi_{M1}=R(i)-p=r(i;A)-\overline{p}+\frac{1}{2}\{[R(i)-C(e)]-[r(i;A)-c(e;B)]\}$$

(2) $$\Pi_{M2}=p-C(e)=\overline{p}-c(e;B)+\frac{1}{2}\{[R(i)-C(e)]-[r(i;A)-c(e;B)]\}$$

Equations (1) and (2) basically indicate the fact that each manager receives her no-trade payoffs plus one half of the gains from trade. The price negotiated at date 2 is determined from the two equations.

De Meza and Lockwood (1998) point out that the assumption regarding the division of *ex post* surplus is not entirely innocuous[15]. First of all, it involves an implicit assumption about the bargaining procedure. Specifically, trades with the spot market are modeled as inside options, that is, if the bargaining procedure is similar to the alternating-offer game of Rubinstein (1982), where each manager receives the no trade payoffs while continuing to bargain with the other manager (and forfeits the no trade payoff once an agreement is reached), then in the limit as discounting tends to zero, the perfect equilibrium payoffs of the two managers will be the ones in equations (1) and (2).

This, however, is not the only way in which the bargaining procedure can be modeled. If instead, bargaining between the two managers and utilization of the spot market are mutually exclusive in the sense that trading in the spot market requires a termination of bargaining, then the no-trade payoffs are outside options and the *ex post* payoffs in this case are different[16]. The difference is not a trivial one, and as De Meza and Lockwood show, some of the results proved in Hart (1995) and Hart and Moore (1990) are no longer true if one alters the bargaining procedure. For example, these papers establish that complementary assets should be owned together, but this is no longer true if no trade payoffs are outside options: non-integration may be optimal even if assets are complementary.

Another feature that may cause some reservation is that the *ex post* bargaining strengths of the two managers remains the same across all ownership structures, irrespective of whether, say, M2 owns a separate firm or if M2 is an employee of M1; in all cases, the gains from trade are divided 50/50. Hart justifies it in the following manner "...This may seem a strong assumption. In fact I would argue that it is a weak assumption. It would be too easy to obtain a theory of the costs and benefits of integration if it were supposed that the bargaining process changes under integration." Ideally, of course, one would like to endogenize the split of the gains from trade and make it sensitive to ownership structures, but this is a complex issue and most of the literature continues to employ an exogenously given division (though not necessarily the 50/50 one) of the gains from trade.

The model then proceeds through backward induction. Once the division of ex post

[15] Chiu (1998) deals with the same issue in the context of Hart and Moore (1990)
[16] See Osborne and Rubinstein (1990) for an analysis of bargaining with outside options.

surplus has been determined for given levels of investments at date 1, the managers choose investment levels non-cooperatively at date 1 to maximize their *ex ante* payoffs, which is given by equations (1) and (2) net of the cost of investments. At date 0 a property right allocation is assigned, that is, the sets A and B (where $A \cup B = \{a1, a2\}$ and $A \cap B = \phi$) are chosen that maximize the total *ex ante* surplus from trade. Hart assumes that the following inequalities hold:

(3) $R'(i) > r'(i; a_1, a_2) \geq r'(i; a_1) \geq r'(i; \phi)$, for all $i \in (0, \infty)$

(4) $|C'(e)| > |c'(e; a1, a2)| \geq |c'(e; a2)| \geq |c'(e; \phi)|$, for all $e \in (0, \infty)$

Given (3) and (4), Hart then shows that one can make the following comparisons for investment levels across different ownership structures:

(5) $i^* > i_1 \geq i_0 \geq i_2$

(6) $e^* > e_2 \geq e_0 \geq e_1$

Here, the superscript '*' indicates the first-best situation where the managers can choose their investments co-operatively, and the subscripts '0', '1', and '2' refer to situations where investments are chosen non-cooperatively (the second best) and the ownership structures are no integration, type 1 integration and type 2 integration, respectively.

There are two points that are worth noting in equations (5) and (6). First, there is 'under-investment' and the second best investment levels are always lower than the first best ones, which is evident from the strict inequalities in (5) and (6). The under-investment result need not always be true, and arises in this instance from assumptions (3) and (4). In Grossman and Hart (1986), for example, either over-investment or under-investment can occur. Secondly, the two equations provide some insight into why property rights matter: they affect the incentives of the managers to invest. To see this, consider a situation where the assets are owned separately (M1 owns a1 and M2 owns a2), so that there are two non-integrated firms. The investment levels of the two managers are i_0 and e_0. Now, if ownership of a2 is given to M1 (type 1 integration), M1's incentives to invest increase $(i_1 \geq i_0)$, whereas M2's incentives to invest fall ($e_1 \leq e_0$). The opposite is true under type 2 integration. Integration thus provides the manager acquiring the asset with higher incentives to invest, but lowers the incentives of the other manager.

At date 0, therefore, ownership of assets is allocated in a manner that maximizes date 1 surplus, that is, the optimization problem is:

$$\underset{k}{Max}\, R(i_k) - C(e_k) - i_k - e_k, \text{ for } k = 0,1,2.$$

Here, it is assumed that the cost of investment can be measured in monetary terms, and can therefore be represented by i_k and e_k, where i_k and e_k are the date 1 non-cooperative choice of investment levels under type k integration.

Finally, at date 1 the division of *ex ante* surplus is determined by the bargaining strengths of the two managers at date 1, which in turn depends on the number of potential trading partners each manager has. So, for example, if M1 is unique and there are a number of alternatives for M2, then M1 has all the *ex ante* bargaining power, and will therefore appropriate all of the date 1 surplus except the reservation utility of M2. Note that the bargaining strengths *ex ante* may be quite different than the strengths *ex post*; once the two agents are locked into a relationship, they make relationship specific investments and this alters their bargaining capabilities.

We conclude this section by examining why the allocation of property rights matters in this model. As the model assumes that date 2 gains from trade are split 50/50 between the two managers, and that date 1 surplus is divided on the basis of the degree of competitiveness, both of which are insensitive to the property right allocation at date 0, it is not entirely apparent at first glance what the repercussions are of different ownership structures. Essentially, there are two effects. First, the ownership structure affects the inside options of the two firms, that is, $r(i;A)$ and $c(e;B)$ in equations (1) and (2) depend on the date 0 allocation of the assets[17]. Even if the division of the gains from trade is insensitive to the asset distribution, the division of the *ex post* surplus when trade occurs, $R(i) - C(e)$, is not: different ownership structures provide the managers with different shares of a given surplus (the share of each manager is different even if the size of the pie is the same). Secondly, since at date 1 the managers' anticipation of their shares is sensitive to the asset distribution, their incentives to invest at date 1 varies with ownership structures as well. This

[17] Note that if we sum up equations (1) and (2), the total is the value of the ex post surplus when trade occurs: $R(i) - C(e)$.

in itself affects the size of the pie, $R(i)-C(e)$, that is available at date 3. At date 0, the ownership structure is chosen that maximizes the size of the pie, net of investment costs.

So, for example, starting off with no integration, if both assets are given to M1, this increases her inside option at date 2, and therefore encourages her to invest more at date 1. The effect on M2 is the opposite: her incentives to invest are reduced. M1's increased investment raises R, whereas M2's decreased investment raises C. The net effect on the total surplus is ambiguous, and depends on whose investment is more 'important' relative to the costs of investment, and ownership of the assets is given to that manager. If both managers make investments that are important, then no integration may be optimal.

The methodology described above is summarized in the following figure, where an arrow indicates that one stage affects another.

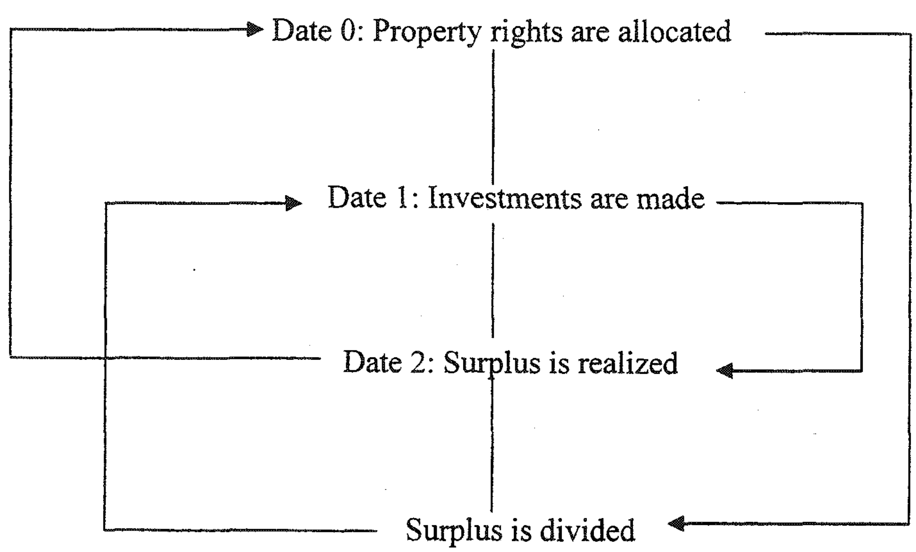

Figure 1.1: The methodology of the property rights theory

Recently, the property rights framework, which bases its analysis on incomplete contracts, has come under scrutiny for lacking rigorous foundations. In particular, papers such as Tirole (1999) and Maskin and Tirole (1999a and 1999b) offer the following criticism against the incomplete contracting methodology. As mentioned before, incomplete contracts are invoked on the basis of the existence of some transaction cost, such as unforeseen contingencies, costs of describing the various contingencies, and the cost of enforcement. At the same time, the incomplete contract literature assumes that agents are rational enough that they can foresee the payoffs associated with their actions, even if they have trouble foreseeing the exact contingencies that may arise. The criticism put forward is that there is a tension between the two; the minimum level of rationality that is guaranteed by the fact that agents can solve a complex dynamic programming problem and the fact that payoffs are foreseeable should enable agents to contract on payoffs, leaving the details of the physical characteristics of trade to be determined once the state of the world is realized. The contract can then be implemented using message games that force agents to truthfully reveal the physical details, making the inability to describe trade in advance irrelevant. In essence, the rationality necessary for dynamic programming makes redundant the very transaction cost arguments that are used to motivate incompleteness of contracts.

The argument is compelling, but rests on the assumption that rational individuals can commit not to renegotiate. Segal (1999) and Hart and Moore (1999) establish models that indicate that if parties cannot commit to not renegotiate the original contract, the optimal contract can often be incomplete. In essence, they argue that describability and transaction costs do play an important role in justifying incomplete contracts. Though a fascinating topic for future research, the models that are developed in the subsequent chapters do not attempt to establish rigorous foundations for incompleteness, and simply invoke transactions cost as a source for incompleteness of contracts. Instead, an attempt is made to extend the basic property rights theory in two directions that have not been addressed by the literature.

2.2 The property rights theory in this dissertation

The power of the incomplete contract approach employed by the property rights theory is evident from the wide range of economic issues that have been analyzed with these tools. Grossman and Hart (1986) provide a more generalized version of the model presented in the previous section to allow for lateral, as well as, vertical integration, whereas Hart and Moore (1990) extend the analysis to multiple agents and multiple assets. Besides the general modeling of integration, the incomplete contracting framework has been applied to various issues, such as: describing the financial structure of firms (Aghion and Bolton (1992), Hart and Moore (1994), Hart (1995, Chapters 5-8)); the ownership of innovation (Aghion and Tirole (1994), Tao and Wu (1997)); the analysis of authority relationships within firm (Aghion and Tirole (1997)); the existence and behavior of multinational enterprises (Chung (2001), Rahman (2001)); land ownership in agriculture (Hueth and Melkonian (2002)); and public versus private ownership (Schmidt (1991, 1996), Laffont and Tirole (1991), Shapiro and Willig (1990))[18].

This dissertation applies the property rights theory to two specific contexts. The first deals with the choice between exports and foreign direct investment facing a firm commencing operations in a foreign country, and the second examines the choice facing firms in the oil industry of whether to co-operate or not while extracting oil. In doing so, an attempt is made to extend the property rights theory in two directions: an examination of the value of information when contracts are incomplete, and the impact of incomplete contracts on situations characterized by non-excludability of assets and rivalry over the ownership of assets.

The dissertation is organized into two papers. The first paper (Chapter II) brings out a somewhat surprising feature of the property rights model: information may lose value when contracts are incomplete. There is a small body of work within contract theory which deals with the issue of whether more information is valuable, and whether agents will choose to remain uninformed (Cremer (1995), Schmidt (1991, 1996), Aghion and Tirole (1997)). These papers show that, in general, a 'principal' may choose to remain uninformed when this provides incentives for an 'agent' to exert more effort. Chapter II shows that when contracts are incomplete, the principal may choose to choose to remain uninformed even in the absence of such strategic considerations. Specifically, the second best nature of the problem when contracts are incomplete may be sufficient to ensure that an economic agent deliberately chooses to remain uniformed.

Chapter II addresses this issue in the context of a firm internationalizing its operations in a foreign country that the firm could be familiar or unfamiliar with. Location specific factors are incorporated that determine the choice between exports and foreign direct investment. Apart from considering similarities in cost structures and asset specificity of

---

[18] This list is by no means exhaustive; Hart (1995) remains a classic reference for a survey of the literature.

investments as location specific factors, the view is presented that the ability to gain familiarity with an initially unfamiliar foreign country can be location specific as well. The paper identifies a 'similarity effect' and a 'familiarity effect' that may reinforce or oppose one another. In the former case, the mode of entry chosen by the firm remains the same irrespective of whether the firm is familiar with the foreign country or not, while in the latter situation the firm may choose different entry modes depending on its familiarity with the foreign market.

The second issue that this dissertation addresses relates to the type of property right arrangements that the literature handles. In particular, most existing papers deal with private property and a few, such as Laffont and Tirole (1991) and Shapiro and Willig (1990), compare private and public ownership. In spite of being considered as polar opposites as far as property right arrangements go, public and private ownership have one feature in common: the state or the private firm has all (or at least a majority of) the residual control rights. A crucial right of an owner of an asset, whether the owner is the state or a private firm, is the ability to exclude individuals from the use of the asset. The state, for example, can decide who can have access to a plant that it owns, and similarly, a private owner of a plant can decide who to employ. Hart and Moore (1990) recognize the right of excludability that private ownership often entails: "We suppose that the sole right possessed by the owner of an asset is his ability to exclude others from the use of that asset. That is, the owner of a machine can decide who can and who cannot work on that machine, the owner of a building can decide who can and who cannot enter the building, the owner of an insurance company's client list can decide who has and who does not have access to the list, and so forth."

There are assets, however, that are owned privately but lack this feature of excludability and this is the focus of the second paper (Chapter III), which analyzes a situation where multiple agents own drilling licenses for common pool of oil. A license to drill for oil provides the owner of the license the right to extract oil, but does not provide her with the right to exclude any other licensee from utilizing the common pool. In such situations, the competitive extraction of oil leads to significant rent dissipation as competing oil firms vie for possession of oil. Numerous authors have suggested that the solution is to co-operate over the extraction of oil; the common property solution often advocated is a unitization agreement. A puzzle confronting the oil industry, however, is that firms are often reluctant to voluntarily enter unitization agreements.

Chapter III looks at the impact of the incomplete nature of the unitization agreement in resolving this. Specifically, the chapter compares two alternative organizational modes for extracting oil from a common pool: competitive extraction and unitization. The literature dealing with this issue usually presumes that unitization is efficient compared to competitive extraction, and looks for frictions in the real world to explain why an inefficient mode (competitive extraction) may be chosen. This chapter questions this common presumption, and argues that in the presence of incomplete contracts, unitization may not always be surplus enhancing relative to competitive extraction. The chapter demonstrates that both competitive extraction and unitization can involve inefficiencies; the inefficiency of the former arises from the common pool problem, while the inefficiency of the latter arises from the fact that oil firms are forced to sign incomplete contracts. Under certain circumstances, unitization may actually involve a lower surplus than competitive extraction. The chapter thus attempts to expand the property rights literature to account for institutions other than private and public ownership, such as common property.

II. FAMILIARITY WITH A FOREIGN MARKET AND THE CHOICE BETWEEN EXPORTS AND FOREIGN DIRECT INVESTMENT

1. Introduction

Of the various alternatives available to a firm while entering a foreign country[1], the choice between exports and foreign direct investment (FDI) has received wide attention in the OLI (ownership-location-internalization) literature of multinational enterprises (MNEs)[2]. The traditional approach has been to focus on the locational factors that affect the decisions of a firm commencing operations abroad. Locational factors encompass features that vary across countries, such as trade barriers, transport costs and factor resources. Factor resources play a crucial role in the development of the theory and, in essence, variations in factor endowments among countries highlight the importance of economic similarities (and differences) in justifying the choice of FDI (and hence the existence of MNEs). An MNE is viewed throughout this paper in a manner similar to Caves' (1996): "an enterprise that controls and manages production establishments—plants—located in at least two countries"[3].

Within the literature addressing this issue, two distinct types of models have emerged. The first deals with horizontal investment[4] between similar countries (Markusen (1984), Brainard (1993), Markusen and Venables (1998)), and involves evaluating the trade off between the added fixed costs of opening a new plant in the foreign country, and the savings from reduced trade costs (tariffs and transport costs). The second type attempts to explain vertical investment (Helpman (1984), Helpman and Krugman (1985)), and analyzes the

[1] See Root (1994) for a fairly exhaustive list of these alternatives.
[2] There has been extensive research on the choice between FDI and other modes of entry, such as franchising, and contract manufacturing. In this paper, we focus attention on the choice between exports and FDI. See Section 6 for a discussion on extensions incorporating other modes of entry.
[3] There are a number of definitions of an MNE in the literature (see Buckley and Casson (1985), for instance). Similarly, there are different interpretations of what constitutes FDI (Buckley and Roberts (1982), for example, outline five features that distinguish FDI from portfolio investments.
[4] Horizontal investment usually refers to production for sale within the host country, while vertical investment conveys the idea of a vertical slicing of the production process across countries with the output being sold to a country other than where the intermediate stage is located. So defined, the distinction between the two can often be vague. For example, horizontal investment may require vertical *integration* within the host country. Indeed, in the model presented in the section 3, this is true. Shatz and Venables (2000) define horizontal FDI as, '...FDI designed to serve local markets...', and vertical FDI as, '...FDI in search of low cost inputs...', thereby tying the definitions to the motivation behind the investment. In this paper, the distinction between the two rests on the market that is intended to be served.

choice of location of a distinct 'headquarters' and production facility. Assuming no frictions to trade, when differences in relative endowments between the host and source countries are sufficiently high, there are incentives for the headquarters and production facility to be located in different countries[5].

Locational advantages can explain the choice between exports and FDI, but cannot explain why a firm prefers FDI to contract manufacturing or technology licensing when entering a foreign country. The advantages of internalization are invoked in the OLI framework to establish this preference (Ethier (1986), Buckley and Casson (1985)). The determinants of the choice between FDI and other modes of entry are not the main focus here, though Section 6 touches briefly upon this matter.

The literature on MNEs has, however, paid scant attention to the effect that familiarity with a foreign market has on the decisions of a firm. With some exceptions[6], typically, the lack of familiarity is viewed simply as an additional fixed cost of FDI[7]. This manner of modeling familiarity with a foreign country provides little insight into the manner in which informational deficiencies can affect decision-making. The importance of familiarity in this paper rests on the notion that the mode of entry can affect the *ability* of a firm to gain familiarity with an unfamiliar market. Since FDI and exports differ in terms of where productive activity is located, it follows that the ability to gather information is location specific as well.

The underlying idea is that FDI forces the home country firm to set up shop within the foreign country. In doing so, the firm has to come in close contact with various economic and social entities within the foreign country, such as the final consumer, the different factors

---

[5] Markusen, Venables, Konan and Zhang (1996) integrate the analysis of both horizontal and vertical investment in a general equilibrium framework, and the results of earlier models emerge as special cases.

[6] Horstmann and Markusen (1996), look at the role of information in a choice between FDI and licensing in a complete contract, hidden information framework. Their analysis emphasizes asymmetric information between the home and foreign country firms, which allows the latter to gather information rents in a licensing contract. In the model presented in this paper, however, information is symmetric when all decisions are made. Moreover, Horstmann and Markusen do not incorporate any location specific features (indeed, the choice between FDI and licensing does not depend on any), and in this sense their model does not have an element of 'internationalism' to it.

[7] Hirsch (1976), for example, discusses the costs of coordinating foreign operations. These costs are likely to be greater in markets that are less familiar. Caves (1996) states that, "The typical entrepreneur, a native of some particular country, brings to his or her business activities...its peculiar 'ways of doing things'... in foreign lands, it must incur a fixed cost of learning how things are done abroad..."

of production, the legal system, the polity, etc. These entities within the foreign market serve as sources of information that aid the home country firm in learning about the foreign market. Exporting, on the other hand, does not involve the same kind of close interaction with agents in the foreign country. To keep the analytics tractable, we consider the extreme case where exports involve no learning, so that learning occurs *only* in the case of FDI[8]. Section 2 takes a closer look at the role of familiarity.

The model developed in Section 3 to analyze the issue of the optimal mode of entry of the home country firm is based on the incomplete contract literature (Grossman and Hart (1986), Hart and Moore (1991), Hart (1995)), which emphasizes that when contracts are incomplete, ownership of assets matter because they confer the owner with residual control rights. The analysis in sections 4 and 5 draws its foundations from Hart (1995), but is primarily concerned with the fact that the presence of incomplete contracts force agents to interact in a second best world, rather than on issues of asset ownership and residual control rights[9].

In a situation where the home country firm is initially unfamiliar with the foreign country, therefore, while making a choice between exports and FDI the firm is, among other things, making the choice of whether to be informed about the foreign country later on (FDI) or not (exports). Section 5 shows that in a world of contractual incompleteness more information can lower the payoff of the firm, leading to the firm to opt for exports over FDI. From a theoretical point of view, the analysis examines the value of information when contracts are incomplete, and adds to the extant literature in contract theory that deals with the issue of whether a decision maker will choose deliberately to remain uninformed (Cremer (1995), Schmidt (1991,1996) Aghion and Tirole (1997)).

In Cremer (1995) the principal can choose between an efficient and inefficient monitoring technology; the efficient technology lowers the principal's cost of determining

[8] This assumption can be dropped to allow for exports to involve learning as well. For example, exporting may allow the manager of the home country firm to learn the relevant information with some probability whereas learning occurs under FDI with certainty, though this adds to the complexity of the analysis without altering the main results. What is important is that FDI is a superior way of learning about the foreign market compared to exports.

[9] The residual control rights play an important role in the choice between alternative modes of entry that locate productive activity within the foreign country; see Section 6. The section incorporates the idea presented in Chung (2001), that these alternative modes of entry differ in terms of asset ownership, and can be evaluated on the basis of managerial incentives to exert effort.

the agent's performance *ex post*. When informed, an *ex ante* threat by the principal to fire an agent on the basis of realizations of output is not credible as the actual performance and ability of the agent can be verified *ex post*. The benefit of choosing the efficient technology (or being informed) is that the principal can gauge the quality of the agent better, while the cost is a weakening of incentives: the removal of the threat of firing forces the principal to pay a higher wage in order to elicit the same effort from the agent. If the latter effect dominates, the principal will choose to remain uninformed later on by selecting the inefficient monitoring technology *ex ante*.

Schmidt (1991,1996) employs a similar intuition in the context of privatization. The choice of the government of whether to privatize an enterprise is linked to the information about a parameter that affects the cost of production; the parameter is the private information of the owner of the firm. Schmidt adopts the idea put forth in Arrow (1975), Riordan (1990) and Shapiro and Willig (1990), among others, that ownership allows access to private information, and that the ability to retrieve some information such as the accounts of a firm are residual rights accruing to the owner of a firm. In such a scenario, the benefit of state ownership is that the information about the parameter allows for *ex post* efficiency in production levels, while the cost is a lower incentive for the manager of the enterprise to exert effort *ex ante*. Under privatization, the asymmetry in information allows the manager/owner to extract an information rent which induces her to exert a higher effort level; when this effect is strong enough the government opts for privatization, even though this lowers the information that the government has about the functioning of the enterprise. Similarly, Aghion and Tirole (1997) show that formal authority provides control over decisions in an organization, but at the same time affects an agent's incentives to communicate information. When there is a tension between the two, the principal may chose to retain control in spite of being less informed.

They key feature in these papers is that the benefit of being less informed for a principal is the positive effect that this has on the incentives of an agent to exert effort. In the model presented here, this is not so: there is no such strategic motivation, and the fact that the principal may chose to remain uninformed stems from the second best nature of the problem when contracts are incomplete. By choosing to remain uninformed (through exports) the principal commits to exerting a higher effort in unfavorable states of the world. This has a positive effect on her *ex ante* expected payoff, and under some circumstances may induce her to opt for exports over FDI.

The only other paper, at least to the knowledge of this writer, which incorporates the property rights theory to analyze the optimal mode of entry into a foreign market is Chung (2001). Chung's main focus, however, is in providing an explanation for the choice of FDI over exports, contract manufacturing and licensing when both the home and foreign countries are developed nations. Specifically, Chung assumes that the two countries are *identical* from the point of view of the cost and technology conditions prevalent for the production of an input required by the MNE. This paper extends Chung's analysis to include dissimilarities between countries, apart from introducing familiarity into the framework.

There are, therefore, two dimensions of the problem that we focus on here; these are identified in Sections 4 and 5 as a 'similarity effect' and a 'familiarity effect' that may reinforce or oppose one another. In the former case, the mode of entry into a foreign country chosen by a firm remains the same irrespective of whether the firm is familiar with the foreign country or not, while in the latter situation, the firm may choose different entry modes depending on its familiarity with the foreign market.

The remaining sections are organized as follows. Section 2 analyzes the issue of familiarity from theoretical and empirical points of view in some depth. Section 3 outlines the basic model of horizontal investment. Section 4 investigates the optimal entry decisions of a firm when entering a country it is familiar with, while Section 5 introduces unfamiliarity into the analysis. Section 6 discusses extensions and concludes.

2. Familiarity with foreign markets

Johanson and Vahlne (1978) hypothesize that FDI is the culmination of a gradual process of learning where an initially naïve internationalizing firm matures into an MNE, and posit that knowledge of foreign markets is conditioned by the 'psychic distance' between countries. The psychic distance (also referred to as 'cultural distance', which is the term adopted in this paper[10]) between two countries is defined as "...the sum of factors preventing the flows of information *from and to* the [foreign] market. Examples are differences in language, education, business practices, culture, industrial development, etc..." (emphasis added). Their hypothesis predicts that the unfamiliarity with a foreign market implies an incremental process of internationalization at two levels: first, a firm is likely to increase its involvement within a particular country in a gradual manner beginning with modes involving low forms of commitment such as exports, and engaging in FDI after it has become familiar with the foreign market; secondly, firms enter familiar markets to commence with and internationalize their operations progressively to less familiar territories. Some empirical studies of one or the other of the two predictions have produced favorable results (Davidson (1980), Johanson and Wiedersheim-Paul (1975), Nicholas (1983)), while others have reported negative findings (Benito and Girsprud (1992), Yu (1990), Turnbull (1987)). At a conceptual level, there has been valid criticism put forward that the predictions of the internationalization hypothesis extends to all firms irrespective of the characteristics of their product and the markets that they serve, thereby ignoring the idiosyncratic features of a particular market and firm level motivations of cost reduction (Reid (1984)).

In spite of its drawbacks, the theory propounded by Johanson and Vahlne provides an interesting way of looking at the notion of familiarity, as being determined by the ease with which information flows between countries. Specifically, a high level of cultural distance serves as a barrier to the flow of information between firms in different countries, and this is the approach adopted here to introduce familiarity into the analysis. The development of the notion of familiarity in this paper rests on two assumptions: first, that cultural distance affects the flow of information *both from and to* foreign markets, and secondly, as mentioned in Section 1, that different modes of entry produce different learning effects for an unfamiliar entrant into the foreign market. The analysis is static in the sense that the mode of servicing the market is a one-time choice; Section 6 discusses briefly how the notion of familiarity affords a dynamic version of the model.

---

[10] Luostarinen (1980) defines cultural distance as "...the sum of factors creating, on the one hand, a need for knowledge, and on the other hand, barriers to knowledge flow and hence also for other flows between the home and target country."

Casual observation seems to suggest that firms are more likely to enter unfamiliar countries through exports rather than FDI. It seems reasonable to posit that, in general, developed countries are more familiar with each other, and are less so with developing countries[11]. Developed countries have comparable per capita incomes, are highly industrialized, have similar technologies, function under democratic political frameworks, have well defined legal systems, share social, cultural and religious ties, have a history of close economic interaction, and so on; all these factors favor increased familiarity between countries. Developing countries often do not share many of these features with developed countries. Most outward and inward flows of FDI occur among the developed countries, and developed countries also account for much of the existing stock of FDI. In 1996, for example, 69 percent of the US FDI stock, 68.4 percent of the Japanese and 64.3 percent of the European were positioned in developed countries (Shatz and Venables (2000) [12]). Similarly, the UNCTAD World Investment Report (2000) indicates that in 1999, developed countries received roughly three fourths of world FDI inflows, and accounted for a little more than 90 percent of outflows.

There are circumstances that promote familiarity between developed and developing countries as well: geographic proximity, language and former imperialist connections come readily to mind as examples. Thus, one finds that Latin America accounts for 19 percent of US stock of FDI in 1996, while Asia accounts for only 8.6 percent, though part of this may be explained by the fact that roughly 20 percent of affiliate production in Latin America is exported to the US, while only 15 percent is imported back from Asia (that is, investment in Latin America is vertical and motivated by lower transport costs associated with geographic proximity rather than by familiarity). Allowing for vertical investment does not, however, explain Japanese trends in FDI. There too we find that geographic proximity plays a role in general FDI patterns: 16.4 percent of Japanese FDI stock in 1996 was positioned in Asia, while 12 percent was positioned in Latin America. But in contrast to the US, a greater percentage of Latin American affiliate production was sold in Japan as compared to Asian affiliate production (28.7 percent versus 18.7). This would imply that the investment in Asia

[11] UNCTAD classifies the European Union, United States, Australia, New Zealand, Canada, Switzerland, Norway, Iceland, Gibraltar, South Africa, Japan and Israel as developed countries.
[12] All data, unless stated otherwise, are from this source.

is largely horizontal (as compared to Latin America), which would follow our general hypothesis that Japan's greater familiarity with South East Asian countries conditions its FDI patterns. Former colonial ties also have influence on FDI distribution. For example, for the period from 1992 to 1994, France's total FDI flows to Morocco totaled \$287 million, while only \$56 million was invested in South Africa. On the other hand, UK's flows to Morocco amounted to \$90 million while that to South Africa was \$1.3 billion.

Geographic proximity, similar language and ethnic ties, in fact, lower the culture distance between any two countries. Intuitively, one would imagine that geographic proximity would favor exports rather than FDI due to low transport costs, but empirical evidence points to the contrary: FDI often occurs between neighboring countries. The familiarity between neighboring countries may serve to reconcile this. Though sparsely researched, there appears to be evidence that developing country MNEs invest in familiar, often developing, countries: "...the host countries tend to be nearby and/or familiar nations: Indian firms to English-speaking and Argentine firms to Spanish-speaking countries." (Caves (1996), pp 241). Wells (1983, pp 78) discusses the role of ethnic ties in influencing FDI decisions, "The extensive overseas communities of Chinese and Indian origin have served to encourage exports that have led to investments. Ethnic ties have also, in some cases, provided a direct link that generated investment without previous export". Even within the developed world, geographic proximity and language seem to play a significant role. Thus, we find that in 1996, UK and Canada held more than 27 percent of US FDI stocks, that two thirds of EU FDI flows between 1992 and 1994 stayed within the EU, and that in 1994 Northern EU and the Nordic countries accounted for 60 percent of total Swedish foreign affiliate production (Ekholm (1998)).

The above discussion illustrates that firms often display a propensity to serve familiar markets through FDI rather than exports. This does not imply, of course, that familiarity with foreign markets is the only motivating factor for FDI decisions. Sections 4 and 5 serve to highlight that a mix of cost, technology and informational considerations determines the choice of entry mode into a foreign country.

3. The model of horizontal investment

Consider a situation where a home country (*HC*) firm, operated by a manager M_1, decides to enter a foreign country (*FC*) in order to sell exactly one unit of a product, *Z*, to the local market[13]. To produce the output, which is assumed to be new to *FC*, for sale to the final market M_1 requires a unit of input, *Q*, and an asset a_1. The asset a_1 is a proprietary, non-human asset owned by the home country firm before it enters *FC*, and could represent, for example, a technology, a brand name, a patent, or even a specific management technique possessed by the home country firm. The asset is transferable across countries at no cost to the home-country firm. We assume that no physical asset (or plant) is required to convert the input into an output.

M_1 has two options open to her. She could either produce *Z* in *HC* and export it to *FC*, or she could enter *FC* through FDI. The first option requires her to purchase the input from a manager in *HC*, M_2, who produces the input *Q* in conjunction with a physical asset a_2. Alternatively, M_1 could produce *Z* in *FC* by procuring the input from a local firm run by a manager M_2^* who produces *Q* using a plant a_2^*. M_1, M_2 and M_2^* are assumed to be risk neutral.

Exports and FDI are characterized not only by the location of production activities, but also by ownership of assets[14]. Specifically, we shall assume that M_1 owns both assets a_1 and a_2 in *HC*[15]. In a similar fashion, FDI entails M_1 acquiring ownership of a_2^* in *FC* from the local firm, and hiring M_2^* as an employee[16]. FDI, therefore, involves M_1 owning of both a_1 and a_2^* in *FC*.

---

[13] The firm is therefore considering horizontal investment.

[14] M_1 is assumed to be wealthy enough to purchase any asset.

[15] This is not an unduly restrictive assumption, though. From the point of view of a theory of MNEs, where by definition an MNE is characterized by ownership of physical assets in at least two countries, it would appear to make sense to allocate ownership of assets in *HC* in this manner (especially as this model *has* only two countries). See Section 6 for a more general analysis where ownership structures in *HC* may be different from those described here.

[16] FDI is therefore through acquisition rather than through green-field investment. If the cost of constructing a new plant is assumed to be the same as the transfer made to the local firm for acquiring a_2^*, the only difference between the two is the presence of M_2^* as an employee when the MNE acquires a_2^*.

(A) Contracts, information and timing

The cultural distance between the two countries, which may be either large or small, is determined exogenously (by nature) at date 0. We assume that all managers know the cultural distance between *HC* and *FC*. At date 1 the home country firm enters *FC*, and has to decide on the mode of entry (Figure 2.1 on page 29 depicts the timeline). Exporting results in M_1 entering into a contractual relationship with M_2, while FDI involves a contractual relationship between M_1 and M_2^*. As in Hart (1995), contracts are incomplete and prescribe only a property right allocation and lump-sum transfers[17]. The *ex ante* bargaining strength is assumed to lie entirely with M_1, and M_2 (M_2^*) has a reservation payoff of $\bar{U}$ ($\bar{U}^*$).

The revenue received from the sale of the output in *FC* depends, among other things, on the 'type' of the market. The high type market has a higher reservation price for *Z*, and values the product sold by the M_1 more than a low type. The parameter $\theta \in \{\theta_H, \theta_L\}$, where $\theta_H > \theta_L$, describes the market type, and is chosen by nature at date 0 along with the cultural distance. Whether a manager learns the value of the parameter, and when, is conditioned by the cultural distance and the physical location of the manager.

The size of the cultural distance influences informational flows both to and from *FC*. The two way flow of information affects, on the one hand, the ability of M_1 and M_2 to gather information about the foreign market, as well as the knowledge that managers in *FC* have about the exact nature of the product *Z*, on the other.

We model the learning process in the following manner. A manager costlessly learns the value of θ when two requirements are satisfied:

(i) The manager has access to information networks within *FC*, and therefore has the ability to collect information within *FC*. This capability is achieved under two circumstances: the manager is physically located within *FC* (the manager is a 'local' manager), or, if the cultural distance is small when the manager is located in *HC*.

(ii) The manager is familiar with the product.

[17] The incompleteness of a contract arises from the inability to describe the exact nature of the input to be traded at a later date. *Ex post* announcement games, as in Maskin (1977), are also precluded by assuming that such mechanisms are too costly to implement.

Regarding the first condition, it is reasonable to posit that information networks are always accessible to local managers, but managers physically located in *HC* have access to these networks only when the cultural distance is small. Similarly, managers in *HC* are familiar with the exact nature of *Z*, but managers in *FC* are familiar with *Z* only when the cultural distance is small[18]. Thus, M_2^* always satisfies the first condition (but not necessarily the second), while M_1 and M_2 always satisfy the second (though not necessarily the first). This implies the following:

(a) When the cultural distance is small, all managers satisfy both requirements, so all managers know the value of θ before date 1. If M_1 decides to enter *FC* through exports, she enters into a contractual relationship with M_2 at date 1, and both managers know the value of the parameter. If M_1 selects FDI as the mode of entry, M_1 and M_2^* enter into a contractual agreement at date 1 with symmetric information about the value of θ.

(b) When the cultural distance is large, neither M_1 nor M_2 satisfies the first requirement before or after date 1[19]. Exports, therefore, involve no learning by M_1 and M_2[20]. With FDI, M_2^* fails to satisfy the second requirement at date 1, and neither M_2^* nor M_1 knows the value of θ at date 1. FDI, however, involves a process of learning for both M_1 and M_2^*. By establishing a physical presence in FC, M_1 becomes a local entity in the foreign country, establishes information networks within the country, and learns the value of θ between dates 1 and 2. Similarly, since M_1 and M_2^* share a close relationship, M_2^* learns the exact nature of the product between dates 1 and 2 (without incurring any cost), and the value of θ is revealed to M_2^* at some time between these dates. The time-line is summarized in Figure 2.1. The arrow between dates indicates that the specified activities occur strictly between those dates.

---

[18] Recall that the product is new to *FC*.

[19] Since exporting does not involve the establishment of a physical presence in *FC*, even after entry into the foreign market, neither manager becomes a 'local' entity, and they fail to satisfy the first condition.

[20] The model ignores the presence of consulting services that can be used by the home country managers to gather information about *FC* when the cultural distance is large. Section 6 discusses this issue.

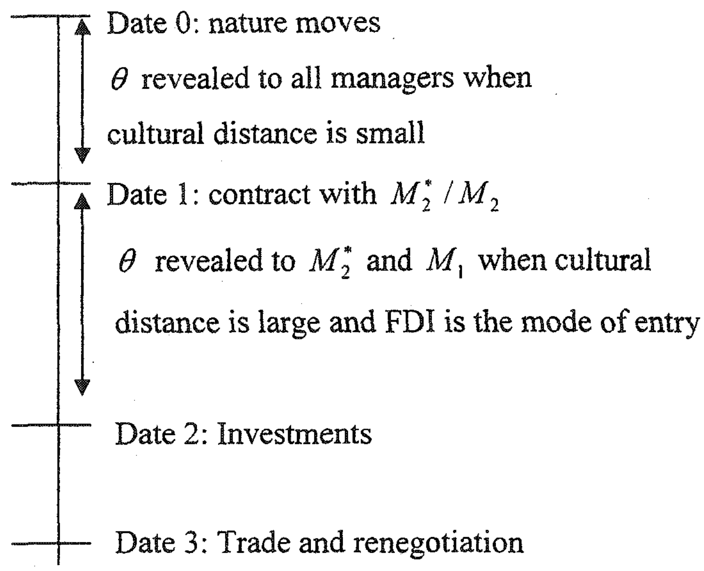

Figure 2.1: The timeline

Definition 1: M_1 (or the home-country firm) is *familiar* with the foreign country if she knows the true value of the parameter while entering the foreign country at date 1, or in other words, if the cultural distance between the home and foreign countries is small.

It is worth emphasizing that, according to Definition 1, when M_1 is familiar with the foreign country all three managers know the value of the parameter at date 1. Conversely, when M_1 is not familiar with *FC* (which corresponds to the situation where the cultural distance is large), no manager knows the value of the parameter at date 1. In this case, it is assumed that the managers have a common prior belief about the value of the parameter. Specifically, they assess a probability of ρ that $\theta = \theta_H$ and of $(1-\rho)$ that $\theta = \theta_L$, where $\rho \in (0,1)$.

All three managers can make non-contractible, relationship specific investments in human capital at date 2. M_1 undertakes an investment $I \in [0, \infty)$, which also represents the cost of investment. Similarly, the level and cost of investments made by M_2 and M_2^* are $e \in [0, \infty)$ and $i \in [0, \infty)$, respectively. I can be interpreted as investments in human capital that increases the productivity of a_1 in FC[21]. For example, I may represent the effort that M_1 expends in increased sales effort, or modifying the product for sale in *FC*. Similarly, e and i may denote M_2 and M_2^*'s investments in human capital in the form of efforts exerted by them in reducing the cost of the input supplied to M_1[22]. At date 3, trade of the input, sale of the output to the market and renegotiation of the surplus takes place. Exporting the good at date 3 involves a tariff, ϕ, and transport costs, τ.

(B) Revenues, costs and trade

MNEs are often considered to bring some technology to the foreign country, which in our model takes the form of the asset a_1. The underlying technology behind the patent allows for the use of a specialized input to produce the output *Z*. The precise knowledge of how to use the specialized input to produce an output lies with M_1. This knowledge may be to some degree, but not entirely, transferable to others, and the presence of M_1 is required for the efficient use of the patent.

With FDI, after the home country firm enters the host market, M_1 forms a relationship with the manager of a local firm, M_2^*, who learns the specifications of the specialized input and acquires knowledge of how to build a specialized input with the asset a_2^*. As with M_1, this knowledge may be transferable in part, but not entirely, to other agents, making M_2^*'s presence a requirement for producing the specialized input efficiently.

The alternative to trading with each other at date 3 is that M_1 and M_2^* utilize the competitive spot market for the purchase and sale of a 'generic' input priced at $\bar{g}^*$. In this paper, the spot market is modeled as an inside option for the managers. The *ex post* (date 3) division of surplus therefore complies with the Nash bargaining solution.

At date 3, M_1 and M_2^* trade in the input at a negotiated price of n^*. After producing

[21] We assume that I does *not* affect the revenues generated through the use of the patent in *HC*.

[22] It is conceivable that M_2 also supplies a unit of input to M_1 for output sold in *HC* as well. In this case, we assume that M_2's efforts in reducing the cost of the input used to produce output supplied to *FC* is technologically independent from that exerted to reduce the cost of input used to produce output supplied to *HC*. Alternatively, we can assume that different managers are responsible for supplying the input to M_1 for output sold in *HC* and *FC*.

the output, M_1 sells it to the local market and earns revenues of $V(I;\theta_k)$, which depends on the type of the market and on M_1's investment[23]. If trade between M_1 and M_2^* breaks down at date 3, M_1 can procure a generic input from the spot market and produce the output for sale to the final market fetching revenues of $v^F(I;\theta_k)$, where lower-case v denotes the absence of M_2^*'s human capital, and the super-script F refers to the use of a_2^* in producing the input (which occurs only under FDI). The cost of producing the specialized input for M_2^* is $C^*(i)$. Similarly, the cost of producing a unit of the generic input is $c^*(i)$, where the lower-case c indicates the absence of M_1's human capital.

Instead of entering through FDI, if M_1 decides to export to the foreign country she has to purchase the input from M_2 by negotiating a purchase price of n at date 3. The output is then exported to the foreign market and sold for $V(I;\theta_k)$, as before. Alternatively, M_1 can use the spot market in the home country for the purchase of a generic input at a price of $\bar{g}$. This results in the output being sold for $v^{\varepsilon}(I;\theta_k)$, where the superscript ε refers to the case of exports and to the use of a_2 in the production process. Exports also involve tariff and transport costs. The cost of producing the specialized input for M_2 is $C(e)$, and the cost of producing a generic input is $c(e)$. In what follows, we assume that the spot markets in *HC* and *FC* are identical, so that $\bar{g}^* = \bar{g}$.

The three assumptions below outline the relationships between the various revenue and cost variables introduced in this section. In the remainder of the paper, the first derivatives with respect to investments are represented using subscripts.

Assumption 1: For any given *I*,

(i) $V_I(I;\theta_H) > V_I(I;\theta_L)$

(ii) $v_I^F(I;\theta_H) > v_I^F(I;\theta_L)$

(iii) $v_I^{\varepsilon}(I;\theta_H) > v_I^{\varepsilon}(I;\theta_L)$

---

[23] Note that M_1 has access to both a_1 and a_2^* if there is trade between M_1 and M_2^*.

Assumption 2: $\forall I, i$ and e

(i) For M_1: $V_I(I;\theta_k) > v_I^F(I;\theta_k)$ and $V_I(I;\theta_k) > v_I^{\varepsilon}(I;\theta_k)$, $\forall \theta_k$

(ii) For M_2^*: $|C_i^*(i)| > |c_i^*(i)|$

(iii) For M_2: $|C_e(e)| > |c_e(e)|$

Assumption 3: $V(0;\theta_L) > 2$, $|C_e(0)| > 2$ and $|C_i^*(0)| > 2$.

Assumption 1 indicates that the high type market (characterized by θ_H) values Z more than the low type market (characterized by θ_L). Assumption 2 captures the benefits to M_1, M_2^* and M_2 from trading in the specialized input[24]. Assumption 3 is a regularity condition that ensures an interior solution.

For given levels of investments, the date 3 gains from trade between M_1 and M_2[25], and between M_1 and M_2^* are represented by (1a) and (1b), respectively:

(1a) $G(I,e;\theta_k) = [V(I,\theta_k) - C(e)] - [v^{\varepsilon}(I;\theta_k) - c(e)]$

(1b) $G^*(I,i;\theta_k) = [V(I,\theta_k) - C^*(i)] - [v^F(I;\theta_k) - c^*(i)]$.

We assume that there are always (positive) gains from trade for all levels of investment. Finally, we suppose for simplicity that the discount rate is zero.

4. The optimal entry mode with familiarity

We begin with an analysis of the situation in which all managers know that $\theta = \theta_k$, where $k \in \{H, L\}$, at date 1. As section 3 indicated, this situation arises when the cultural distance between *HC* and *FC* is small. Since all managers know the value of the parameter when the MNE enters *FC*, this section isolates the effects that similarities and differences between the two countries have on the choice of the optimal mode of entry. Subsequently, we shall label this the 'similarity effect'.

---

[24] These are monotonicity assumptions made in Hart (1995, pp 37).

[25] Note that since identical tariffs and transport costs are present irrespective of whether the final output was produced using a generic or specialized input, these cancel out in the expression for gains from trade.

4.1 Exporting from the home country

Since the payoff from exports serves as M_1's reservation payoff while deciding whether to establish a physical presence *FC* through FDI, we commence with the description of the outcomes from exports. Consider, first, the situation where M_1 and M_2 can cooperate over the choice of *I* and *e*, or in other words, where efforts are contractible. This represents the first best situation and involves choosing investments at date 2 that maximize the surplus at date 3, that is,

(2) $$\max_{I,e} V(I;\theta_k) - C(e) - I - e - \phi - \tau$$

The first order conditions when $\theta = \theta_k$ are:

(2a) $$V_I(\hat{I}_k^\varepsilon;\theta_k) = 1$$

(2b) $$\left|C_e(\hat{e}_k)\right| = 1$$

Here the '^' denotes familiarity with *FC*, and 'ε' denotes exports. The total surplus in this case is given by:

(2c) $$\hat{S}_k^\varepsilon = V(\hat{I}_k^\varepsilon;\theta_k) - C(\hat{e}_k) - \hat{I}_k^\varepsilon - \hat{e}_k - \phi - \tau$$

It can be seen that the maximization problem above is independent of ownership structures. If investments are non-contractible, however, then ownership of assets *does* matter, and we enter the second-best world. Ownership of assets confers a manager with residual control over the use of the asset, which affects the inside options of the managers at date 3, and this in turn has a feedback effect on the managers' incentives to invest at date 2[26]. Given our assumption that the *ex post* gains from trade are split through Nash bargaining, the payoffs to M_1 and M_2 at date 3 are:

$$M_1: \quad v^\varepsilon(I;\theta_k) - \bar{g} - \phi - \tau + \tfrac{1}{2}\{[V(I;\theta_k) - C(e)] - [v^\varepsilon(x,I) - c(e)]\}$$

$$M_2: \quad \bar{g} - c(e) + \tfrac{1}{2}\{[V(I;\theta_k) - C(e)] - [v^\varepsilon(I;\theta_k) - c(e)]\}$$

The price of the input negotiated at date 3 will be:

$$n(I,e;\theta_k) = \bar{g} + \tfrac{1}{2}[V(I;\theta_k) - v^\varepsilon(I;\theta_k)] - \tfrac{1}{2}[c(e) - C(e)].$$

At date 2, M_1 and M_2 will choose investments non-cooperatively to maximize the payoffs

[26] See Grossman and Hart (1986) or Hart (1996) for a more detailed exposition.

from trade at date 3. The maximization problem for M_1 and M_2 are:

$$M_1: \quad \max_I v^{\varepsilon}(I;\theta_k) - \bar{g} + \tfrac{1}{2}G(I,e;\theta_k) - I - \phi - \tau$$

$$M_2: \quad \max_e \bar{g} - c(e) + \tfrac{1}{2}G(I,e;\theta_k) - e$$

The first order conditions are:

$$\text{(3a)} \quad \tfrac{1}{2}V_I(\hat{I}^{\varepsilon}_{k1};\theta_k) + \tfrac{1}{2}v^{\varepsilon}(\hat{I}^{\varepsilon}_{k1};\theta_k) = 1$$

$$\text{(3b)} \quad \tfrac{1}{2}\left|C_e(\hat{e}_{k1})\right| + \tfrac{1}{2}\left|c_e(\hat{e}_{k1})\right| = 1$$

Here, $\hat{I}^{\varepsilon}_{k1}$ and $\hat{e}_{k1}$ denote the optimal second best level of investment for M_1 and M_2, respectively, given the value of the parameter[27].

The total surplus from exporting is:

$$\text{(3c)} \quad \hat{S}^{\varepsilon}_{k1} = V(\hat{I}^{\varepsilon}_{k1}) - C(\hat{e}_{k1}) - \hat{I}^{\varepsilon}_{k1} - \hat{e}_{k1} - \phi - \tau$$

The transfers set in the contract will be such that M_2 will be left with her reservation payoff. The exact magnitude of the transfer at date 1 is sensitive to the 'continuation game' after date 1 and depends on M_2's contribution (in the sense of value added) to the expected surplus. The transfer will therefore entail a payment by M_2 to M_1 of:

$$\text{(3d)} \quad \hat{t}^{\varepsilon}_{k1} = n(\hat{I}^{\varepsilon}_{k1}, \hat{e}_{k1};\theta_k) - C(\hat{e}_{k1}) - \hat{e}_{k1} - \overline{U}$$

M_1's payoff from exporting (when investments are non-contractible) is:

$$\text{(3e)} \quad \hat{\pi}^{\varepsilon}_{k1} = V(\hat{I}^{\varepsilon}_{k1}) - C(\hat{e}_{k1}) - \hat{I}^{\varepsilon}_{k1} - \hat{e}_{k1} - \phi - \tau - \overline{U}$$

We assume that $\hat{\pi}^{\varepsilon}_{k1} > 0$, so that exports are always a profitable proposition for M_1.

4.2 Establishing a physical presence in the foreign country

The alternative that M_1 has to exporting Z to the foreign country is to produce the output in FC after procuring the input in the host country itself[28]. With FDI, this involves purchasing a_2^* from M_2^*, retaining the local manager as an employee, and producing the

[27] The sub-script '1' refers to the fact that the ownership pattern considered in this paper corresponds to the situation that Hart (1995) refers to as 'type-1 integration': M_1 owns both assets.

[28] We do not consider situations where M_1 imports the input into the foreign country in order to produce the output.

input with the local firm as an affiliate of the MNE. If M_1 and M_2^* make contractible investments, they maximize the total value of the expected surplus through their choice of investments and the ownership structure does not matter. The problem, in this case, is:

$$\max_{I,i} V(I;\theta_k) - C^*(i) - I - i$$

The first order conditions resulting from this optimization problem are:

(4a) $V_I(\hat{I}_k^F;\theta_k) = 1$

(4b) $\left|C_i^*(\hat{i}_k)\right| = 1$

The total value of the surplus at date 2 is:

(4c) $V(\hat{I}_k^F;\theta_k) - C^*(\hat{i}_k) - \hat{I}_k^F - \hat{i}_k$

If, on the other hand, investments are non-contractible M_1 and M_2^* maximize, respectively, (5a) and (5b) below:

(5a) $\max_{I} v^F(I;\theta_k) - \overline{g}^* + \frac{1}{2}G^*(I,i;\theta_k) - I$

(5b) $\max_{i} \overline{g}^* - c^*(i) + \frac{1}{2}G^*(I,i;\theta_k) - i$

At date 3, they renegotiate the price of the input, which is given by:

(5c) $n^*(I,i;\theta_k) = \overline{g}^* + \frac{1}{2}[V(I;\theta_k) - v^F(I;\theta_k)] - \frac{1}{2}[c^*(i) - C^*(i)]$

The first order conditions for (5a) and (5b) are:

(6a) $\frac{1}{2}V_I(\hat{I}_{k1}^F;\theta_k) + \frac{1}{2}v_I^F(\hat{I}_{k1}^F;\theta_k) = 1$

(6b) $\frac{1}{2}\left|C_i^*(\hat{i}_{k1})\right| + \frac{1}{2}\left|c_i^*(\hat{i}_{k1})\right| = 1$

With FDI, the total surplus (with optimal levels of investments) is:

(6c) $\hat{S}_{k1}^F = V(\hat{I}_{k1}^F;\theta_k) - C^*(\hat{i}_{k1}) - \hat{I}_{k1}^F - \hat{i}_{k1}$

M_1 will set date 1 transfers in the contract that leave M_2^* with her reservation payoff:

(6d) $\hat{t}_{k1}^F = n^*(\hat{I}_{k1}^F, \hat{i}_{k1};\theta_k) - C^*(\hat{i}_{k1}) - \hat{i}_{k1} - \overline{U}^*$

M_1's expected payoff at date 1 with FDI is, therefore:

(6e) $\hat{\pi}_{k1}^F = V(\hat{I}_{k1}^F;\theta_k) - C^*(\hat{i}_{k1}) - \hat{I}_{k1}^F - \hat{i}_{k1} - \overline{U}^*$

4.3 The optimal mode of entry

We are now in a position to bring together the results of the analysis of the previous sub-sections. Lemma 1 below presents a summary of how the investments levels compare in the various scenarios described before.

Lemma 1: Given Assumptions 1 and 2 if $\theta = \theta_k$ is observed by all the managers at date 1, then:

(i) $\hat{I}_H^{\varepsilon} > \hat{I}_L^{\varepsilon}$ and $\hat{I}_{H1}^{\varepsilon} > \hat{I}_{L1}^{\varepsilon}$

(ii) $\hat{I}_H^F > \hat{I}_L^F$ and $\hat{I}_{H1}^F > \hat{I}_{L1}^F$

(iii) $\hat{I}_k^F > \hat{I}_{k1}^F$ and $\hat{I}_k^{\varepsilon} > \hat{I}_{k1}^{\varepsilon}$, for all $k \in \{H, L\}$

(iv) $\hat{I}_k^{\varepsilon} = \hat{I}_k^F$, for all $k \in \{H, L\}$

(v) $\hat{e}_H = \hat{e}_L$ and $\hat{e}_{H1} = \hat{e}_{L1}$

(vi) $\hat{i}_H = \hat{i}_L$ and $\hat{i}_{H1} = \hat{i}_{L1}$

(vii) $\hat{i}_k > \hat{i}_{k1}$ and $\hat{e}_k^{\varepsilon} > \hat{e}_{k1}^{\varepsilon}$, for all $k \in \{H, L\}$.

Proof: See Appendix.

The intuition behind the first two parts of the lemma is straightforward: Assumption 1 implies that the marginal benefit of investment is greater when the value of the parameter is higher, so M_1 always invests more when $\theta = \theta_H$. Parts (iii) and (vii) are the well-known under-investment results associated with the second-best nature of the problem when contracts are incomplete, and that follow from Assumption 2. Part (iv) states that when investments are contractible, since the marginal incentives to invest for M_1 are identical under exports and FDI for a given value of the parameter, her investments in both cases are identical. Parts (v) and (vi) imply that the incentives to invest for M_2 and M_2^* are independent of θ as their cost functions are not affected by the value of the parameter. This property allows us to simplify the notation slightly, and in the subsequent analysis, we shall use the following: $\hat{e}_H = \hat{e}_L = \hat{e}$, $\hat{e}_{H1} = \hat{e}_{L1} = \hat{e}_1$, $\hat{i}_H = \hat{i}_L = \hat{i}$ and $\hat{i}_{H1} = \hat{i}_{L1} = \hat{i}_1$.

Though Lemma 1 allowed for a comparison of M_1's investment levels across exports and FDI when investments were contractible, in order to facilitate this comparison when investments are non-contractible we have to say something more about the assets a_2 and a_2^*. This is achieved with the next definition.

Definition 2: $\forall k \in \{H, L\}$ and $\forall I$

(i) The assets a_2 and a_2^* are *similar assets for* M_1*'s inside option* if:

$v_I^{\varepsilon}(I;\theta_k) = v_I^F(I;\theta_k)$

(ii) The asset a_2 is defined to be *superior (inferior) for* M_1*'s inside option* if:

$v_I^{\varepsilon}(I;\theta_k) \geq (\leq) v_I^F(I;\theta_k)$, with the strict inequality holding for some *I*.

Definition 2 describes how the assets a_2 and a_2^* may differ in terms of affecting M_1's inside option, and provide an idea of how specific M_1's investments are to the two assets. If the two assets are similar, then M_1's investments are equally well suited to both assets. If, on the other hand, M_1's investment is more specific to a_2 than to a_2^*, then a_2 is superior to a_2^* for M_1's inside option. We assume, however, that M_1's marginal benefit from investment within a relationship with another manager does *not* depend on the asset that was used to produce the input. This implies that $V_I(I;\theta_k)$ is the same whether a_2 or a_2^* was used to produce the specialized input.

Lemma 2: Given Assumption 2, $\forall k \in \{H, L\}$,

(i) $\hat{I}_{k1}^{\varepsilon} = \hat{I}_{k1}^F$, if a_2 and a_2^* are similar assets for M_1's inside option.

(ii) $\hat{I}_{k1}^{\varepsilon} \geq \hat{I}_{k1}^F$, if a_2 is superior to a_2^* for M_1's inside option.

(iii) $\hat{I}_{k1}^{\varepsilon} \leq \hat{I}_{k1}^F$, if a_2 is inferior to a_2^* for M_1's inside option.

Proof: Follows from comparing (3a) and (6a) using Definition 2.

Lemmas 1 and 2 do not tell us how the investment levels of M_2 compare with those of M_2^*, a comparison which is needed in order to evaluate exports versus FDI. Definition 3 below makes such a comparison possible.

Definition 3: $\forall i = e$

(i) *HC* and *FC* are defined as having *similar costs* if:

$C(e) = C^*(i)$, $|C_e(e)| = |C_i^*(i)|$ and $|c_e(e)| = |c_i^*(i)|$

(ii) *HC* is defined as having a *cost advantage (disadvantage)* with respect to *FC* if:

$C(e) \leq (\geq) C^*(i)$, with the strict inequality holding for some $i = e$.

$|C_e(e)| \geq (\leq) |C_i^*(i)|$, with the strict inequality holding for some $i = e$, and,

$|c_e(e)| \geq (\leq) |c_i^*(i)|$, with the strict inequality holding for some $i = e$.

FC has a cost advantage (disadvantage) if the inequalities are reversed.

Definition 3 compares the cost conditions prevalent in the two countries. The cost disadvantage (or advantage) of the domestic country over the foreign country can arise for a number of reasons. For example, consider the situation where the production of the input requires a third factor, say (unskilled) labor, apart from the physical asset (plant) and the managers' efforts. Even if the plants and managers in the two countries are identical, it may be that the foreign country has more, and hence cheaper, unskilled labor than the domestic country, giving the foreign country a cost advantage (in absolute and marginal sense). This can be thought of as a case where the cost *functions* are different. Of course, it can also be possible that the plants themselves are different in terms of technology, which may reinforce or counteract the first advantage (and if the cost functions are the same, cause it to shift). In any case, Definition 3 represents the final situation after all the possible effects are taken into account.

<u>Lemma 3</u>: $\forall k \in \{H, L\}$,

(a) If investments are chosen cooperatively, then:

(i) $\hat{e} = \hat{i}$, when *HC* and *FC* have similar costs.

(ii) $\hat{e} \geq \hat{i}$, when *HC* has a cost advantage.

(iii) $\hat{e} \leq \hat{i}$, when *HC* has a cost disadvantage.

(b) If investments are non-contractible, then:

(i) $\hat{e}_1 = \hat{i}_1$, when *HC* and *FC* have similar costs.

(ii) $\hat{e}_1 \geq \hat{i}_1$, when *HC* has a cost advantage.

(iii) $\hat{e}_1 \leq \hat{i}_1$, when *HC* has a cost disadvantage.

Proof: Using Definition 3, we get (a) by comparing (2b) and (4b), and (b) by comparing (3b) and (6b).

<u>Proposition 1</u>: Suppose that investments are non-contractible and that $\overline{U} = \overline{U}^*$. If M_1 is familiar with the foreign country, a_2 and a_2^* are similar for M_1's inside option, and the domestic and foreign countries have similar costs, then FDI yields a higher payoff for M_1 than exports for any positive value of tariffs and transportation costs.

Proof: See Appendix.

The intuition behind Proposition 1 is that if the domestic and foreign countries are identical, M_1 can always 'replicate' the domestic country ownership structure in the foreign country, and save on the tariff and transport costs associated with exporting. Since the ownership structure in the home country is that M_1 owns both assets, replicating that ownership structure in the foreign country would involve owning both assets there, which is precisely the ownership structure under FDI. Proposition 1 is the result obtained in Chung (2001), which is interpreted in Chung's paper as a framework to account for FDI between two developed countries[29]. Exports between developed countries, however, are possible only if tariffs and transport costs are zero, and even then, only weakly so. This sub-optimality of

[29] Though the notion of familiarity of the foreign country is absent in Chung's analysis.

exports result is unattractive as it runs counter to the empirical observations that exports between developed countries, and even two way intra-industry exports, are high. The following proposition identifies situations where exports can be optimal.

Proposition 2: Suppose that investments are non-contractible, M_1 is familiar with the foreign country, and that $\overline{U} = \overline{U}^*$. The optimal mode of entry is described in Table 1 below for all cases not covered in Proposition 1 (which are represented by '-'). In the table 'E' represents exports; more specific bounds on the terms 'high' and 'low' in the table are described in the proof.

Table 2.1: The optimal mode of entry with familiarity

	'Low' tariffs and transport costs			'High' tariffs and transport costs		
	Asset a_2 is:			Asset a_2 is:		
	Similar	Superior	Inferior	Similar	Superior	Inferior
Costs are similar	-	E	FDI	-	FDI	FDI
HC has cost advantage	E	E	E/FDI	FDI	FDI	FDI
HC has cost disadvantage	FDI	E/FDI	FDI	FDI	FDI	FDI

Proof: See Appendix.

Proposition 2 states in formal terms the fairly obvious idea that for exports to emerge as the optimal mode of entry, there must be some advantage to locating the production process in *HC*[30], and tariffs and transport costs must be sufficiently low. This advantage may emerge due to two reasons. First, the plant used in *HC* may be more specific to M_1's investments than the plant in *FC*. Secondly, a cost advantage in *HC* permits M_2 to invest more compared to M_2^*, *ceteris paribus*, and the surplus is higher under exports unless transport costs and tariffs erode this advantage. Situations where one country has an

[30] This is the idea behind the locational advantage in the OLI framework as well.

advantage in one of these factors and a disadvantage in the other can result in FDI or exports, depending on the magnitude of the effects.

Regarding the first reason, when a_2 is superior for M_1's inside option, M_1's *ex post* share of the surplus rises, *ceteris paribus*. This encourages M_1 to invest more under exports than under a comparable ownership structure (FDI) in *FC*. Exports dominate as a mode of entry when the advantage from the plant in *HC* being more specific to M_1's investments is not lost due to high tariffs and transport costs. Though there is nothing unexpected about Proposition 2, it is a useful result given that both exports and FDI are commonly observed in the interaction between countries, especially in the developed world. The analysis suggests that even in situations where input costs are comparable between two countries, firms in industries where significant relationship specificity with upstream assets exist are more likely to serve a familiar foreign market through exports.

5. The optimal mode of entry with unfamiliarity

The previous section isolated the effects that similarities and dissimilarities between countries have on the choice between exports and FDI. In this next section we isolate the effect of familiarity with *FC* by assuming that costs between the two countries are similar, and that assets are similar for M_1's inside option. From Proposition 1, we know that in this case FDI would be chosen as the optimal mode of entry.

When the cultural distance is large, the timing of events described in Section 3 implies that no manager knows the value of the parameter at date 1 when M_1 enters FC. At date 1, therefore, there is symmetric information, and the managers share the common belief that $\theta = \theta_H$ with probability ρ and that $\theta = \theta_L$ with probability $(1-\rho)$. If M_1 enters *FC* through FDI, at some time between dates 1 and 2, the value of θ is revealed to both M_1 and M_2^*. On the other hand, if M_1 decides to service the foreign country market through exports, the large cultural distance prevents both M_1 and M_2 from acquiring information about θ (before and after date 1). The choice of location of productive activity, therefore, affects the knowledge of the parameter that M_1 has at date 2 while making her investment, and the ability to gather information is location specific.

5.1 Exporting when the foreign country is unfamiliar

As before, we commence with the payoffs to M_1 from exports. When investments are non-cooperatively chosen, the maximization problem for M_1 at date 2 is:

$$\max_I \rho v^\varepsilon(I;\theta_H)+(1-\rho)v^\varepsilon(I;\theta_L)-\overline{g}+\tfrac{1}{2}E_\rho[G(I,e;\theta_k)]-I$$

The date 2 expectations of the gains from trade are taken with respect to prior beliefs over parameter values. The first order condition for M_1's problem is:

(7a) $$\tfrac{\rho}{2}[V_I(\widetilde{I}_1^\varepsilon;\theta_H)+v_I^\varepsilon(\widetilde{I}_1^\varepsilon;\theta_H)]+\tfrac{1-\rho}{2}[V_I(\widetilde{I}_1^\varepsilon;\theta_L)+v_I^\varepsilon(\widetilde{I}_1^\varepsilon;\theta_L)]=1$$

The superscript '~' denotes the unfamiliarity of *FC*. For M_2, the problem is:

$$\max_e \overline{g}-c(e)+\tfrac{1}{2}E_\rho[G(I,e;\theta_k)]-e$$

The optimal investment level for M_2 is given by:

(7b) $$\tfrac{1}{2}\left|C_e(\widetilde{e}_1)\right|+\tfrac{1}{2}\left|c_e(\widetilde{e}_1)\right|=1$$

Comparing (7b) with (3b), it is evident that $\widetilde{e}_1=\hat{e}_1$; M_2's investments are not affected by uncertainty about the parameter value. On the other hand, comparing (7a) with (3a), we see that $\hat{I}_{H1}^\varepsilon>\widetilde{I}_1^\varepsilon>\hat{I}_{L1}^\varepsilon$. Intuitively, since M_1 is unfamiliar with *FC* and $\rho\in(0,1)$, she hedges her bets by investing somewhere between the levels that she would have done were she to know with certainty the value of the parameter. Thus, there is an incentive for M_1 to invest more than the full information case when the parameter value is low, and to invest less when the parameter value is high. In the former case, as long as $\widetilde{I}_1^\varepsilon<\hat{I}_L^\varepsilon$, the added effort by M_1 has a beneficial effect in terms of increasing the surplus when the parameter value is actually low, while if she exerts too much effort and $\widetilde{I}_1^\varepsilon>\hat{I}_L^\varepsilon$, the surplus when $\theta=\theta_L$ may be lower than that achieved if she had exerted the effort $\hat{I}_{L1}^\varepsilon$. When $\theta=\theta_H$, the under-investment problem is exacerbated due to unfamiliarity. The expected total surplus at date 1 from exports is:

(7c) $$\widetilde{S}_1^\varepsilon=\rho V(\widetilde{I}_1^\varepsilon;\theta_H)+(1-\rho)V(\widetilde{I}_1^\varepsilon;\theta_L)-C(\hat{e}_1)-\widetilde{I}_1^\varepsilon-\hat{e}_1-\phi-\tau$$

The date 1 contract specifies transfers that leaves M_1 with an expected payoff of:

(7d) $\tilde{\pi}_1^{\varepsilon} = \rho V(\tilde{I}_1^{\varepsilon};\theta_H) + (1-\rho)V(\tilde{I}_1^{\varepsilon};\theta_L) - C(\hat{e}_1) - \tilde{I}_1^{\varepsilon} - \hat{e}_1 - \phi - \tau - \overline{U}$

It is evident that the payoff to M_1 through exporting is different from that under the case where she is familiar with *FC* (see (3c)). In fact, it can be shown that the payoff lies somewhere in between those when M_1 knows whether $\theta = \theta_H$ or $\theta = \theta_L$.

5.2 FDI when the foreign country is unfamiliar

When M_1 assesses her payoffs while establishing a physical presence in *FC*, she now has to take into account the learning possibilities that arise with FDI. Specifically, M_1 and M_2^* will learn the value of the parameter between dates 1 and 2, and the date 2 situation is identical to that in Section 4. At date 1, however, neither M_1 nor M_2^* know the value of θ, and so their belief about the expected value of surplus is:

(8) $\tilde{S}_1^F = \rho\hat{S}_{H1}^F + (1-\rho)\hat{S}_{L1}^F$

$= \rho[V(\hat{I}_{H1}^F;\theta_H) - \hat{I}_{H1}^F] + (1-\rho)[V(\hat{I}_{L1}^F;\theta_L) - \hat{I}_{H1}^F] - C^*(\hat{i}_1) - \hat{i}_1$

The date 1 contract specifies transfers such that M_2^* is left with her reservation utility, and M_1's payoff will, hence, be:

(9) $\tilde{\pi}_1^F = \tilde{S}_1^F - \overline{U}^*$

Given that costs are similar in the two countries, it follows that $\hat{i}_1 = \hat{e}_1 = \tilde{e}_1$ and that $C^*(\hat{i}_1) = C(\hat{e}_1) = C(\tilde{e}_1)$. Intuitively, the roles of the local and home country manager supplying the input play no role in the comparison of FDI and exports when costs are similar. We can simplify the subsequent analysis by considering what the outcome would have been if (hypothetically) FDI, like exports, involved no learning. It is apparent that M_1's investment must satisfy the first order condition:

(10) $\frac{\rho}{2}[V_I(\tilde{I}_1^F;\theta_H) + v_I^F(\tilde{I}_1^F;\theta_H)] + \frac{1-\rho}{2}[V_I(\tilde{I}_1^F;\theta_L) + v_I^F(\tilde{I}_1^F;\theta_L)] = 1$

Comparing (7a) with (10), we see that $\tilde{I}_1^{\varepsilon} = \tilde{I}_1^F$. Thus, from the point of view of M_1's investments, we can treat the export case as equivalent to the FDI case where M_1 does not learn the value of the parameter (that is, *where* M_1 makes her investment is irrelevant when the assets are similar for M_1's inside option).

Proposition 1 indicated that when the home and foreign countries were economically similar, exports can never emerge as the optimal mode of entry for any positive value of transport costs and tariffs. The question that arises now is whether the same outcome (superiority of FDI) holds when M_1 is unfamiliar with the foreign country at the time of entry. The answer to this question is addressed in Proposition 4, and involves a comparison of $\tilde{\pi}_1^{\varepsilon}$ with $\tilde{\pi}_1^F$. As a result of cost and asset similarity, evaluating the relative magnitudes of $\tilde{\pi}_1^{\varepsilon}$ and $\tilde{\pi}_1^F$ amounts to comparing (11a) and (11b) below:

(11a) $\Pi_{FDI} = \rho[V(\hat{I}_{H1}^F;\theta_H) - \hat{I}_{H1}^F] + (1-\rho)[V(\hat{I}_{L1}^F;\theta_L) - \hat{I}_{L1}^F] - \overline{U}^*$

(11b) $\Pi_E = \rho V(\tilde{I}_1^F;\theta_H) + (1-\rho)V(\tilde{I}_1^F;\theta_L) - \tilde{I}_1^F - \phi - \tau - \overline{U}$

Before we make that comparison, however, it is worthwhile taking a look at what happens when investments can be chosen co-operatively. This is summarized in Proposition 3.

Proposition 3: Suppose that M_1 is unfamiliar with the foreign country, a_2 and a_2^* are similar for M_1's inside option, the domestic and foreign countries have similar costs, and that $\overline{U} = \overline{U}^*$. If managers can cooperate over the choice of investments at date 2, then $\forall \rho \in (0,1)$ exports yield a lower payoff to M_1 than FDI when transport costs and tariffs are non-negative; the results of Proposition 1 hold.
Proof: See Appendix.

In establishing the counterpart to Proposition 3 for the case where investments are non-contractible, our task will be made easier by assuming that the revenue functions are quadratic in *I*. Though this is not necessary for the proposition to hold, the assumption drastically reduces the complexity of the analysis and provides a clearer intuition of the issues involved.

<u>Assumption 4</u>: $V(I;\theta_k)$, $v^F(I;\theta_k)$ and $v_I^{\varepsilon}(I;\theta_k)$ are quadratic functions of *I*.

When $V(I;\theta_k)$, $v^F(I;\theta_k)$ and $v_I^{\varepsilon}(I;\theta_k)$ are quadratic functions of I, the expected marginal revenues at date 2 are linear in I and can be written as, $\forall I$:

$$V_I(I;\theta_H) = a_1 + b_1 I \text{ and } V_I(I;\theta_L) = c_1 + d_1 I$$

$$v_I^F(I;\theta_H) = a_2 + b_2 I \text{ and } v_I^F(I;\theta_L) = c_2 + d_2 I$$

There are two points that are worth noting about these expected marginal revenues. First of all, the marginal revenues from the inside options are expressed in terms of those under FDI; as a consequence of the discussion related to equation (10), these are the only ones we need to consider. Secondly, any (or all) of $a_1, b_1, a_2, b_2, c_1, c_2, d_1, d_2$ can be functions of the parameter. For the remainder of the section, we let $a = a_1 + a_2$, $b = b_1 + b_2$, $c = c_1 + c_2$, and $d = d_1 + d_2$. Note that b and d represent the responsiveness of marginal investments in the second best case to a small increment in investments when the parameter value is high and low, respectively.

Proposition 4: Suppose that investments are non-contractible and that $\overline{U} = \overline{U}^*$, and that Assumption 4 holds. If M_1 is unfamiliar with the foreign country, a_2 and a_2^* are similar for M_1's inside option, and the domestic and foreign countries have similar costs, then exports can yield a higher payoff to M_1 than FDI for *some* belief over the parameter values when tariffs and transportation costs are small enough.

Proof: See Appendix.

Corollary: Two separate sufficient (but not necessary) conditions for exports to yield a higher payoff than FDI for some belief over parameter values in the simple quadratic case when transportation costs and tariffs are small enough are:

$$\text{(12a)} \quad \frac{b}{d} V_I(\hat{I}_{L1}^F;\theta_L) > \frac{V(\hat{I}_{H1}^F;\theta_H) - V(\hat{I}_{L1}^F;\theta_H)}{\hat{I}_{H1}^F - \hat{I}_{L1}^F} + (\frac{b}{d} - 1)$$

$$\text{(12b)} \quad \frac{d}{b} V_I(\hat{I}_{H1}^F;\theta_H) < \frac{V(\hat{I}_{H1}^F;\theta_L) - V(\hat{I}_{L1}^F;\theta_L)}{\hat{I}_{H1}^F - \hat{I}_{L1}^F} + (\frac{d}{b} - 1)$$

Proof: See proof of Proposition 4 in Appendix.

Proposition 4 implies that even if *HC* and *FC* are economically very similar as in Proposition 1, but the foreign country is now an unfamiliar one, exports may emerge as the optimal entry mode. In effect, Proposition 4 isolates the 'familiarity' effect: *ceteris paribus* the lack of familiarity with a foreign country can sometimes (not necessarily always) favor exports as opposed to FDI. The corollary to Proposition 4 outlines sufficient conditions that would guarantee that exports are preferred to FDI (given that assets and costs are similar) when transport costs and tariffs are small enough. In the special case where $b = d$ (12a) reduces to:

$$(13) \quad V_I(\hat{I}_{L1};\theta_L) > \frac{V(\hat{I}_{H1};\theta_H) - V(\hat{I}_{L1};\theta_H)}{\hat{I}_{H1} - \hat{I}_{L1}}$$

This implies that a sufficient condition for Π_E to exceed Π_{FDI} (ignoring transportation costs and tariffs) is for the marginal revenue when $\theta = \theta_L$ evaluated at $\hat{I}_{L1}$ to be greater than the average increment in revenues due to increasing investment from $\hat{I}_{L1}$ to $\hat{I}_{H1}$ in the high state. Expression (12a) conveys the same information, except for adjusting marginal and average increments in revenue for differences in responsiveness of marginal investment to parameter values. (12b) establishes an alternative criterion for enabling exports to yield a higher payoff than FDI.

As the introductory section pointed out, informational asymmetries may allow an agent to gather some informational rent that is detrimental to the principal, but in turn may induce the former to exert a higher effort, which the principal finds beneficial. In the presence of these opposing forces the principal may choose an ownership structure that leaves her uninformed. The interesting feature of Proposition 4, from a more theoretical point of view, is that it occurs entirely due to the second-best nature of the problem, and not due to an informational asymmetry. The learning process between dates 1 and 2 ensures that the 'principal', M_1, is perfectly informed under FDI while making her investment. The fact that the *anticipation* of bargaining at date 3 distorts M_1's incentives to invest at date 2 is the driving force behind Proposition 4 can be readily observed by comparing the results in Propositions 3 and 4.

The distortion in this model takes the form of M_1 choosing to under-invest when contracts are incomplete (Lemma 1). The uninformed investment level lies somewhere in-between the full information levels when the parameter value is high and when it is low. Hence, there is a greater amount of under-investment when M_1 is uninformed and the true value of the parameter is $\theta = \theta_H$ than when she is informed. When $\theta = \theta_L$, however, since M_1's investment level is higher than in the informed case, one of three things may happen: her investment may be lower, equal to or greater than the first best investment level. In the first two cases, if the added benefits on the surplus outweigh the negative effects of an increased under-investment when $\theta = \theta_H$, M_1 may choose to remain uninformed. In the third situation, when uninformed, M_1 over-invests when $\theta = \theta_L$ and under-invests when $\theta = \theta_H$, with the magnitude of the under-investment in the second instance being, again, higher than that when she is informed about the parameter value. In this situation it is still possible depending on the magnitudes of over and under-investments, that exports yield a higher payoff than FDI, especially if the extent of over-investment is small.

In the end, there are two things that determine the circumstances under which Proposition 4 holds: the beliefs over the parameter values and the particular functional form of the revenue functions. Indeed, there can be instances where it depends only on the functional form, and exports may emerge superior to FDI for *any* belief of parameter values, as the following example indicates.

Example: Suppose that for $k \in \{H, L\}$:

$$V(I;\theta_k) = (\alpha + g\theta_k)I - \frac{I^2}{2} \text{ and } v^F(I;\theta_k) = (\beta + u\theta_k)I - \frac{I^2}{2}$$

Suppose that α, β, g and u are such that Assumptions 1 and 2 are satisfied, and that transport and tariff costs are zero. The marginal revenue functions are:

$$V_I(I;\theta_k) = (\alpha + g\theta_k) - I, \text{ and } v_I^F = (\beta + u\theta_k) - I$$

Note that this falls into special case described earlier where b = d. Solving for the second best investment levels when *FC* is a familiar country, we get $\hat{I}_{H1}^F - \hat{I}_{L1}^F = \frac{1}{2}(g+u)(\theta_H - \theta_L)$.

It also follows that $\tilde{I}_1^F = \frac{1}{2}[(\alpha+\beta)+(g+u)(\rho\theta_H+(1-\rho)\theta_L)]-1 = \rho\hat{I}_{H1}^F+(1-\rho)\hat{I}_{L1}^F$. It is easy to verify that the sufficient condition implied by (12a) reduces to $u > 3g$. Therefore, we have that for this example, $u > 3g$ is sufficient for M_1's payoff from exports to exceed the payoff from FDI for *some* belief of the parameter values, as long as transport costs and tariffs are small enough. If we actually go ahead and calculate Π_E and Π_{FDI} (and assume that reservation payoffs for M_2 and M_2^* are equal), we can show that:

$$\Pi_{FDI} = \Pi_E + \tfrac{1}{8}(u+g)\,\mathrm{var}(\theta)(3g-u)+\phi+\tau$$

Here, $\mathrm{var}(\theta) = E(\theta_k^2)-[E(\theta_k)]^2$, with expectations taken over beliefs. This confirms that when $u > 3g$, $\Pi_E > \Pi_{FDI}$ as long as transport costs and tariffs are small enough. In fact, it indicates that this condition is necessary for $\Pi_E > \Pi_{FDI}$ to be true, and whenever it holds $\Pi_E > \Pi_{FDI}$ is true for *any* belief about the parameter values that the MNE may have while entering *FC* (when transport and tariff costs are low enough).

Some idea into why this happens in this example can be gathered from the fact that if we evaluate condition (12b), we end up with the same criterion for the superiority of exports over FDI: $u > 3g$. In effect, since $b = d$, $\frac{\partial \tilde{I}^F(\rho)}{\partial \rho}$ is independent of ρ, and hence conditions (12a) and (12b) do not depend on ρ either. In general, of course, this need not be true.

5.3 Summary: The familiarity and similarity effects

The analysis in this section indicates that when contracts are incomplete, familiarity (or its lack thereof) plays an important role in determining the choice between FDI and exports. The discussion in Sections 4 and 5 identifies two effects at work: a 'similarity effect', which is determined by conditions of cost, technology and asset specificity of investments in *HC* and *FC*, and a 'familiarity effect' that follows from the level of information that the home country firm has about the foreign country. These effects may either counteract or reinforce each other. When these effects reinforce each other, then the results in Table 1 hold when the foreign country is an unfamiliar one as well. On the other hand, when they oppose one another, exports can emerge as the optimal mode of entry when the foreign country is unfamiliar under circumstances where FDI is the optimal entry mode were *FC* to be a familiar country.

If we return to Table 1 and consider the case, for example, where costs are similar but a_2 is inferior for M_1's inside option, the similarity effect by itself favors FDI; the familiarity effect may oppose this when the conditions of Proposition 4 hold and if its effect on M_1's payoffs is large enough, can result in exports being chosen over FDI. Though not derived explicitly here, the tension between the two effects modifies Table 1 to allow for exports to emerge as the optimal mode of entry where FDI would otherwise have been chosen. This fact can help explain some of the evidence presented in Section 2, which indicated that a firm is more likely to enter a foreign country with exports when it is unfamiliar with the foreign country.

6. Extensions and conclusions

6.1 Alternative modes of entry

Though the primary focus of this paper has been the choice between exports and FDI, the analysis can be extended in a fairly straightforward manner to incorporate other modes of entry. The key is to recognize that any ownership pattern over assets a_1 and a_2 that exists in *HC* can be replicated when the home country firm (F_1) enters *FC*. In the previous sections, we assumed that F_1 owned both assets in *HC*; the parallel ownership structure in *FC* corresponds to F_1 owning a_1 and a_2^*. By acquiring a plant (a_2^*) in *FC*, F_1 engages in FDI, ends up owning a plant in both countries and achieves the status of an MNE. Consider an alternative ownership structure in *HC* with F_1 owning a_1 and the downstream firm (F_2) owning a_2. This ownership structure can be replicated in *FC* by allowing the local firm (F_2^*) to retain ownership of a_2^*.

In general, we can define an ownership structure over the assets a_1 and a_2 in *HC* as the ordered pair (A, B)[31], where A is the set of assets owned by F_1 in *HC* and B is the set of assets owned by F_2. Similarly, the ownership structure over a_1 and a_2^* in *FC* is represented by (A^*, B^*), where A^* is the set of assets owned by F_1 in FC and B^* is the set of assets owned by M_2^*.

There are three 'leading' ownership structures (A^*, B^*) that are possible once the MNE has entered the foreign country (the terms in the brackets are those employed by Hart (1995); also see Chung (2001)):

1. *Contract manufacturing* (No integration): F_1 and F_2^* remain independent firms and M_1 contracts for the procurement of Q with M_2^*. Thus, $(A_0^*, B_0^*) = (\{a_1\}, \{a_2^*\})$.
2. *Foreign direct investment* (Type 1 integration): F_1 takes over the physical asset a_2^*, and so ends up owning both a_1 and a_2^*; F_2^* is a subsidiary of F_1. Here, $(A_1^*, B_1^*) = (\{a_1, a_2^*\}, \phi)$.
3. *Franchising* (Type 2 integration): Here F_1 licenses out the use of the patent to F_2^* for a fee. F_2^* takes ownership of the patent, and $(A_2^*, B_2^*) = (\phi, \{a_1, a_2^*\})$.

In considering only the leading ownership structures, we ignore the implausible situation where F_1 owns a_2^* and F_2^* owns a_1 (this pattern is ruled out in *HC* as well). If we also assume that F_2 does not own any assets in *FC* and that F_2^* does not own any assets in *HC*, we arrive at the following convenient definition.

Definition 4: The two countries have the *same ownership structure* if F_1 owns the same number of assets in both countries. With some abuse of notation, this situation can denoted by $(A, B) = (A^*, B^*)$[32].

[31] Note that the ownership structure (A, B) must satisfy the condition that $A \cap B = \phi$ and $A \cup B = \{a_1, a_2\}$. Similarly, (A^*, B^*) must satisfy $A^* \cap B^* = \phi$ and $A^* \cup B^* = \{a_1, a_2^*\}$. These conditions imply that joint ownership of assets is precluded and that any given asset must be owned by some manager.

[32] The definition implies that if, for example, $(A, B) = (\{a_1, a_2\}, \phi)$, then $(A, B) = (A^*, B^*)$ if and only if $(A^*, B^*) = (\{a_1, a_2^*\}, \phi)$.

If we suppose that contract manufacturing and franchising also allow for learning, the analysis of Sections 4 and 5 are straightforward to duplicate. The choice between the alternative modes of entry will be conditioned upon the following factors.

(a) For a given (A, B)[33], the choice between exports and the same ownership structure in *FC* is determined by the location specific features described in Sections 4 and 5: tariffs, transport costs, the similarity effect and the familiarity effect.

(b) The choice between FDI, contract manufacturing and franchising is determined on the basis of managerial incentives to exert effort (Hart (1995))[34].

In terms of the alternatives considered, the analysis in this sub-section is similar to Chung (2001). Unlike Chung's paper, however, exports can emerge as the optimal mode of entry. The necessary condition for this to happen is that the similarity and familiarity effects result in exports being chosen in (a). If, in addition, the choice in (b) is the same ownership structure as the *HC*, exports will be chosen as the mode of entry when transport costs and tariffs are low.

6.2 The value of consulting services

Sections 5 showed that when contracts are incomplete, agents may deliberately choose to remain uninformed. In other words, more information is not always valuable when contracts are incomplete, and may actually cause a lowering of surplus. The model can be extended to analyze the value of consulting services. To see this, consider, in general, the interaction between two managers M_1 and M_2 involved in a vertical or lateral relationship characterized by incomplete contracts. Suppose that their knowledge of the trading environment, while symmetric, is incomplete. If a consulting agency could be hired to gather the information *after* the contractual relationship has been established[35], the question that

---

[33] We are therefore considering a partial analysis since we do not allow for ownership patterns in *FC* to affect the ownership structure in *HC*. One would, in general, expect that MNEs optimize over ownership of assets across countries so that ownership structures in the foreign country may well have implications for the assets that the MNE decides to retain in its home country.

[34] If these modes involve different potential for learning, this will play a role in the choice.

[35] We assume here that the consulting service is hired by both managers, so that any information gathered by the service is shared by the managers, and information continues to be symmetric. We also assume that any information provided by the agency is hard, (in the terminology of Aghion and Tirole (1997)), which implies that all information can be verified by the managers.

arises is: would they choose to hire these services, even if it were free?

Section 5 indicates that the answer may well be no. In the context of exports in our model, suppose that M_1 and M_2 could commission a market survey at or after date 1 (the date at which they enter a relationship) when the cultural distance is large[36], so that information about the parameter is revealed between dates 1 and 2. The export situation is now identical to FDI, and the result in Proposition 4 holds: exporting without subsequently commissioning the survey may yield a higher surplus than exports with additional knowledge about the foreign country gathered after entry.

Clearly the same result would hold if the survey was commissioned, and the information arrives, before date 1. From the perspective of the time at which the market research is commissioned, uncertainty regarding the parameter value will ensure that the result of Proposition 4 holds while deciding whether to export with or without information about the parameter. It is worth noting, however, that the very notion of a large cultural distance implies that effective consulting services are unavailable before date 1. If such services were available to begin with, they would ensure the flow of information between the two countries, which contradicts the logic of a large cultural distance. In essence, in the context of multinational enterprises, when the cultural distance is large consulting agencies may be constrained by the same features that prevent the managers from gathering information: inexperience with the foreign country or the lack of familiarity with the product.

In a more general context, the model suggests that market research in the presence of incomplete contracts may not always be surplus enhancing, even if costless and guaranteed to reveal information about market.

6.3 Other extensions

The model developed in Section 3 was concerned primarily with horizontal investment, that is, with the MNE desiring to serve the host country market. Though horizontal investment accounts for much of the world FDI activity, vertical investment is an

---

[36] Recall that gathering information requires the home country managers to establish a physical presence in the foreign country when the cultural distance is large. Since the home country managers do not establish a physical presence with exports, the value of the consulting services arises from creating a relationship with an entity that has access to information networks within the foreign country.

important component of the FDI flows into developing countries, and sometimes into developed countries as well (Canada, for example, receives substantial vertical investment from US MNEs). The model in this paper can be modified to account for this phenomenon. A straightforward extension would be to simply let $V(I;\theta)$ denote the revenues gathered by the home country firm from sale of the output in HC, where θ now represents the parameter in *HC*. In keeping with the logic of the model, M_1 in this case knows the value of the parameter while M_2^* may not (depending on the cultural distance). In this circumstance one may find asymmetric information at date 1.

The model presented in this paper was a static one, in the sense that the MNE's mode of servicing the market was a one-time choice. In reality, one finds that firms often choose a mode of entry (such as exports) while entering the foreign market, but decide to switch to another mode (say FDI) after some length of time[37]. Though the dynamic aspect of the foreign investment process is the central feature of the internationalization hypothesis, little work has been done to formalize this (the internationalization hypothesis literature does not rely much on analytical rigor), and as Buckley and Casson (1981) state, "Analyses of the optimal timing of foreign dynamic investment decisions have been curiously lacking in the general literature of multinational enterprises." The context of MNEs also brings out the need for a more dynamic perspective of the integration process than the one envisaged by the incomplete contract literature, which views integration as a static phenomenon.

A simple (though possibly uninteresting) way to incorporate a switch in the mode of servicing the market without changing the structure of the model would be to allow for the home country firm to re-evaluate its options at date 2. This is an unattractive proposition in the model as it stands because of the assumption that exports preclude any learning. A more meaningful way to allow for a switch in property rights, though far more complicated, is to repeat the game while introducing a stochastic variable and dynamic linkages (say the first stage effort affects the second stage outcome as well). This allows for exports to have some learning benefits as well: the outcome of the stochastic variable in the first stage serves as a signal for M_1, resulting in the updating of initial beliefs.

[37] There are instances of firms disinvesting as well, which cannot be captured by the model as it stands.

6.4 Conclusions

This paper has attempted to assess the value of information when contracts are incomplete in the context of the choice of entry mode by a firm into a foreign country. While the traditional OLI theory of multinational enterprises has been correct in emphasizing the cost minimization motives of MNEs that are dictated by economic similarities and dissimilarities between countries, the literature on MNEs has not given due recognition to the fact that the impact of familiarity with the foreign country extends beyond fixed organizational costs. In reality, familiarity with a foreign country *and* economic similarities appear to play a central role in the decisions of an MNE. This was summarized in our stylized model by isolating a familiarity effect and a similarity effect. When these work in opposite directions, a firm may choose to enter an unfamiliar foreign country with exports rather than FDI under the same circumstances where it would have chosen to enter with the latter had the country been a familiar one. From a theoretical point of view, the paper showed that more information may not always be of value when contracts are incomplete, and agents may choose to remain uninformed. This happens even if information is symmetric, and players are not driven by strategic considerations regarding the reaction of other players to the possession of more or less information.

Appendix

Proof of Lemma 1:

(i) The first inequality follows from comparing equation (2a) for the cases where k=H and k=L while using (1a), and the second form comparing equation (3a) for k=H and k=L, using (1a) and (1c).

(ii) Similar to the manner in part (i), the first inequality is derived from (4a) using (1a), and the second from (6a) using (1a) and (1b).

(iii) Given Assumption 2, this is derived for a given k from (4a) and (6a) using for FDI, and from (2a) and (3a) for exports.

(iv) Follows by comparing (2a) and (4a).

(v) The first equality follows from (2b) and the second from (3b).

(vi) The first equality is obtained from (4b) and the second from (6b).

(vii) This follows from Assumption 2 and the equations (4b) and (6b) for M_2^* and equations (2b) and (3b) for M_2.

■

Proof of Proposition 1:

The payoff to M_1 from exporting is, for $k \in \{H, L\}$:

(3e) $\hat{\pi}_{k1}^{\varepsilon} = V(\hat{I}_{k1}^{\varepsilon}; \theta_k) - C(\hat{e}_1) - \hat{I}_{k1}^{\varepsilon} - \hat{e}_1 - \phi - \tau - \overline{U}$.

The payoff to M_1 from FDI is:

(6e) $\hat{\pi}_{k1}^{F} = V(\hat{I}_{k1}^{F}; \theta_k) - C^*(\hat{i}_1) - \hat{I}_{k1} - \hat{i}_1 - \overline{U}^*$.

Lemma 2 indicates that if the two assets a_2 and a_2^* are similar, then $\hat{I}_{k1}^{F} = \hat{I}_{k1}^{\varepsilon}$, while if the costs of producing the input are also similar, then lemma 3 shows that $\hat{e}_1 = \hat{i}_1$. Substituting these values in (6e) and employing definition 3 to note that the costs are the same, the payoff to M_1 from FDI, given that $\overline{U} = \overline{U}^*$ is:

$$\hat{\pi}_{k1}^{F} = V(\hat{I}_{k1}^{\varepsilon}; \theta_k) - C(\hat{e}_1) - \hat{I}_{k1}^{\varepsilon} - \hat{e}_1 - \overline{U} = \hat{\pi}_{k1}^{\varepsilon} + \phi + \tau$$

We see that the payoff to M_1 from a presence in the foreign country is greater than that from exports when $\phi + \tau > 0$. This establishes the proposition.

■

Proof of Proposition 2:

The various possibilities in Table 1 are dealt with case by case.

<u>Case 1</u>: Costs are similar but a_2 is superior to a_2^* for M_1's inside option.

From Lemma 2 (ii) we can see that $\hat{I}_{k1}^{\varepsilon} \geq \hat{I}_{k1}^{F}$. Since the costs in the two countries are similar, Lemma 3b (i) indicates that $\hat{e}_1 = \hat{i}_1$, and from definition 3, the cost functions are the same as well. This, along with the fact that $\overline{U} = \overline{U}^*$ implies that the payoff to M_1 from exporting and FDI are given by (3e) and (6e), which after making substitutions for similarity in the costs yield:

(I) $\hat{\pi}_{k1}^{\varepsilon} = V(\hat{I}_{k1}^{\varepsilon};\theta_k) - C(\hat{e}_1) - \hat{I}_{k1}^{\varepsilon} - \hat{e}_1 - \phi - \tau - \overline{U}$

(II) $\hat{\pi}_{k1}^{F} = V(\hat{I}_{k1}^{F};\theta_k) - C(\hat{e}_1) - \hat{I}_{k1}^{F} - \hat{e}_1 - \overline{U}$

M_1's payoff under FDI from any level of investment by M_1 and M_2^* is given by (noting the similarity of costs and equality of reservation payoffs):

$$V(I;\theta_k) - C(i) - I - i - \overline{U}\,.$$

Taking the derivative of this with respect to I and evaluating at $\hat{I}_{k1}^{F}$, we get that:

(III) $V_I(\hat{I}_{k1}^{F};\theta_k) - 1 > \tfrac{1}{2}V_I(\hat{I}_{k1}^{F};\theta_k) + \tfrac{1}{2}v_I^F(\hat{I}_{k1}^{F};\theta_k) - 1 = 0$

The first inequality follows from Assumption 2, and the equality from (6a). (III) indicates that increasing M_1's investment in (II) raises M_1's payoff. Consider, first, the case when transport and tariff costs are zero in (I) above. Since $\hat{I}_{k1}^{\varepsilon} \geq \hat{I}_{k1}^{F}$, using (III) to compare (I) and (II) implies that $\hat{\pi}_{k1}^{\varepsilon} \geq \hat{\pi}_{k1}^{F}$. So if tariffs and transport costs are zero, exports yield M_1 at least as high a payoff as FDI, which is the first part of the proposition. When tariffs and transport costs are not zero, then exporting is optimal if:

$$V(\hat{I}_{k1}^{\varepsilon};\theta_k) - C(\hat{e}_1) - \hat{I}_{k1}^{\varepsilon} - \hat{e}_1 - \phi - \tau - \overline{U} \geq V(\hat{I}_{k1}^{F};\theta_k) - C(\hat{e}_1) - \hat{I}_{k1}^{F} - \hat{e}_1 - \overline{U}$$

That is, if:

$$\phi + \tau \leq [V(\hat{I}_{k1}^{\varepsilon};\theta_k) - \hat{I}_{k1}^{\varepsilon}] - [V(\hat{I}_{k1}^{F};\theta_k) - \hat{I}_{k1}^{F}]$$

When this inequality holds, tariffs and transport costs are low enough for exports to emerge as the optimal mode of entry. If this inequality, 'high' tariff and transport costs will ensure that FDI is chosen as the mode of entry.

Case 2: Costs are similar but a_2 is inferior to a_2^* for M_1's inside option.

The payoffs to M_1 from exporting and from FDI are given by (I) and (II) above.

Given that a_2 is inferior to a_2^* for M_1's inside option, lemma 2 (iii) implies that $\hat{I}_{k1}^{\varepsilon} \leq \hat{I}_{k1}^{F}$. From (III), we have that increasing M_1's investment in (II) raises M_1's payoff, which is exactly what exporting does.

In other words, for any $\phi + \tau \geq 0$:

$$\hat{\pi}_{k1}^{F} = V(\hat{I}_{k1};\theta_k) - C(\hat{e}_1) - \hat{I}_{k1} - \hat{e}_1 - \overline{U}$$

$$\geq V(\hat{I}_{k1}^{\varepsilon};\theta_k) - C(\hat{e}_1) - \hat{I}_{k1}^{\varepsilon} - \hat{e}_1 - \overline{U} \geq V(\hat{I}_{k1}^{\varepsilon};\theta_k) - C(\hat{e}_1) - \hat{I}_{k1}^{\varepsilon} - \hat{e}_1 - \phi - \tau - \overline{U} = \hat{\pi}_{k1}^{\varepsilon}$$

The first follows from (III) and the fact that $\hat{I}_{k1}^{\varepsilon} \leq \hat{I}_{k1}^{F}$ and the reason for the first inequality is obvious.

Case 3: a_2 is similar to a_2^* for M_1's inside option, and *HC* has a cost advantage. Lemma 2 (i) implies that $\hat{I}_{k1}^{\varepsilon} = \hat{I}_{k1}^{F}$ and Lemma 3b (ii) that $\hat{e}_1 \geq \hat{i}_1$. Since $\overline{U} = \overline{U}^*$, M_1's payoff from exporting and FDI are:

(IV) $\quad \hat{\pi}_{k1}^{\varepsilon} = V(\hat{I}_{k1}^{\varepsilon};\theta_k) - C(\hat{e}_1) - \hat{I}_{k1}^{\varepsilon} - \hat{e}_1 - \phi - \tau - \overline{U}$

(V) $\quad \hat{\pi}_{k1}^{F} = V(\hat{I}_{kA}^{\varepsilon};\theta_k) - C^*(\hat{i}_1) - \hat{I}_{k1}^{\varepsilon} - \hat{i}_1 - \overline{U}$

M_1's payoff from any level of investment by M_1 and M_2^* under FDI is given by:

$$V(I;\theta_k) - C^*(i) - I - i - \overline{U}$$

Taking the derivative with respect to i, and evaluating at $\hat{i}_1$, we get:

(VI) $\quad \left|C_i^*(\hat{i}_{\hat{B}^*})\right| - 1 > \frac{1}{2}\left|C_i^*(\hat{i}_1)\right| + \frac{1}{2}\left|c_i^*(\hat{i}_1)\right| - 1 = 0$

The first inequality can be deduced from Assumption 2 and the second equality follows from (6b). The surplus in the foreign country therefore rises when the investment of the manager supplying the input increases, and falls when it decreases. When $\phi + \tau = 0$, we have that:

$$\hat{\pi}_{k1}^{\varepsilon} = V(\hat{I}_{k1}^{\varepsilon};\theta_k) - C(\hat{e}_1^{\varepsilon}) - \hat{I}_{k1}^{\varepsilon} - \hat{e}_1 - \overline{U}$$

$$\geq V(\hat{I}_{kA}^{\varepsilon};\theta_k) - C^*(\hat{e}_1) - \hat{I}_{k1}^{\varepsilon} - \hat{e}_1 - \overline{U} \geq V(\hat{I}_{k1}^{\varepsilon};\theta_k) - C^*(\hat{i}_1) - \hat{I}_{k1}^{\varepsilon} - \hat{i}_1 - \overline{U} = \hat{\pi}_{k1}^{F}$$

Here, the first inequality comes from the definition of a cost advantage, and the second from (III) and the fact that $\hat{e}_1 \geq \hat{i}_1$.

Even if $\phi + \tau > 0$ exports can still yield at least as high a payoff to M_1 if:

(VII) $\phi + \tau \leq [C^*(\hat{i}_1) - \hat{i}_1] - [C(\hat{e}_1) - \hat{e}_1]$

Sufficiently low tariff and transport costs as in (V) will ensure the optimality of exports; the failure of the inequality will result in FDI.

Case 4: a_2 is similar to a_2^* for M_1's inside option, and *HC* has a cost disadvantage.

The payoffs to M_1 from exporting and FDI are given by (IV) and (V) above. Lemma 2 (i) implies that $\hat{I}_{k1}^{\varepsilon} = \hat{I}_{k1}^{F}$ and Lemma 3b (iii) that $\hat{e}_1 \leq \hat{i}_1$. Now, we have that for $\phi + \tau \geq 0$:

$$\begin{aligned}
\hat{\pi}_{k1}^{F} &= V(\hat{I}_{k1}^{F};\theta_k) - C^*(\hat{i}_1) - \hat{I}_{k1}^{F} - \hat{i}_1 - \overline{U} \\
&= V(\hat{I}_{k1}^{\varepsilon};\theta_k) - C^*(\hat{i}_1) - \hat{I}_{k1}^{\varepsilon} - \hat{i}_1 - \overline{U}^* && \text{(Lemma 2 (i))} \\
&\geq V(\hat{I}_{k1}^{\varepsilon};\theta_k) - C^*(\hat{e}_1) - \hat{I}_{k1}^{\varepsilon} - \hat{e}_1 - \overline{U}^* && \text{(from (IV))} \\
&\geq V(\hat{I}_{k1}^{\varepsilon};\theta_k) - C(\hat{e}_1) - \hat{I}_{k1}^{\varepsilon} - \hat{e}_1 - \overline{U}^* && \text{(cost disadvantage)} \\
&\geq V(\hat{I}_{k1}^{\varepsilon};\theta_k) - C(\hat{e}_1) - \hat{I}_{k1}^{\varepsilon} - \hat{e}_1 - \phi - \tau - \overline{U}^* \\
&= \hat{\pi}_{k1}^{\varepsilon}
\end{aligned}$$

Case 5: a_2 is superior to a_2^* for M_1's inside option, and *HC* has a cost advantage.

From cases 1 and 3, it follows that exports will emerge as the mode of entry if transport and tariff costs are sufficiently low; otherwise FDI will emerge as the mode of entry.

Case 6: a_2 is inferior to a_2^* for M_1's inside option, and *HC* has a cost disadvantage.

From cases 2 and 4, it follows that FDI is the optimal mode of entry.

Case 7: If one country has a cost advantage and the other has a superior asset for M_1's inside option, then either cases 1 and 4 hold, or cases 2 and 3 hold. In that case, it is apparent that which mode of entry is optimal depends on the magnitudes of the opposing forces, and the size of transport and tariff costs.

■

Proof of Proposition 3:

When costs are similar, $\tilde{e}_1 = \tilde{i}_1$. In a manner similar to equations (11a) and (11b) in order to evaluate exports and FDI we need to compare (assuming zero transport costs and tariffs and that $\overline{U} = \overline{U}^*$):

(I) $\Pi_{FDI} = \rho[V(\hat{I}_H^F;\theta_H) - \hat{I}_H^F] + (1-\rho)[V(\hat{I}_L^F;\theta_L) - \hat{I}_H^F]$

(II) $\Pi_E = \rho V(\tilde{I}^\varepsilon;\theta_H) + (1-\rho)V(\tilde{I}^\varepsilon;\theta_L) - \tilde{I}^F$

$\tilde{I}^\varepsilon$ satisfies the first order condition:

(III) $\rho V_I(\tilde{I}^\varepsilon;\theta_H) + (1-\rho)V_I(\tilde{I}^\varepsilon;\theta_L) = 1$

$\hat{I}_H^F$ and $\hat{I}_L^F$, of course, satisfy equation (4a). Comparing (III) with (4a), it is apparent that $\tilde{I}^\varepsilon = \hat{I}_L^F$ when $\rho = 0$ and $\tilde{I}^\varepsilon = \hat{I}_H^F$ when $\rho = 1$, implying that $\Pi_{FDI} = \Pi_E$ when $\rho = 0$ and when $\rho = 1$. Differentiating (II) with respect to ρ:

(IV) $$\frac{\partial \Pi_E(\rho)}{\partial \rho} = V(\tilde{I}^\varepsilon;\theta_H) - V(\tilde{I}^\varepsilon;\theta_L) + [\rho V_I(\tilde{I}^\varepsilon;\theta_H) + (1-\rho)V_I(\tilde{I}^\varepsilon;\theta_L)]\frac{\partial \tilde{I}^\varepsilon}{\partial \rho} - \frac{\partial \tilde{I}^\varepsilon}{\partial \rho}$$

After using (III) to eliminate terms in (IV) and differentiating again, we get that:

(V) $$\frac{\partial^2 \Pi_E(\rho)}{\partial \rho^2} = [V_I(\tilde{I}^\varepsilon;\theta_H) - V_I(\tilde{I}^\varepsilon;\theta_L)]\frac{\partial \tilde{I}^\varepsilon}{\partial \rho}$$

Differentiating (III) with respect to ρ:

$$\frac{\partial \tilde{I}^\varepsilon}{\partial \rho} = -\frac{V_I(\tilde{I}^\varepsilon;\theta_H) - V_I(\tilde{I}^\varepsilon;\theta_L)}{\rho V_{II}(\tilde{I}^\varepsilon;\theta_H) + (1-\rho)V_{II}(\tilde{I}^\varepsilon;\theta_L)} > 0$$

This implies that Π_E is strictly convex in ρ. Since that $\Pi_{FDI} = \Pi_E$ when $\rho = 0$ and when $\rho = 1$ and Π_E is strictly convex in ρ, it follows that $\Pi_{FDI} > \Pi_E$ for all $\rho \in (0,1)$ even when transport costs and tariffs are zero; the inequality will obviously still hold when they are positive. This proves the proposition.

■

Proof of Proposition 4 and Corollary:

To prove Proposition 4, we need to compare (11a) and (11b), given that $\overline{U}^* = \overline{U}$. Suppose that transport and tariff costs are zero; comparing (11a) with (11b) then reduces to comparing:

(I) $\Pi_{FDI} = \rho[V(\hat{I}^F_{H1};\theta_H) - \hat{I}^F_{H1}] + (1-\rho)[V(\hat{I}^F_{L1};\theta_L) - \hat{I}^F_{L1}]$

(II) $\Pi_E = \rho V(\tilde{I}_1^F;\theta_H) + (1-\rho)V(\tilde{I}_1^F;\theta_L) - \tilde{I}_1^F$

Recall that $\forall I$:

$V_I(I;\theta_H) = a_1 + b_1 I$; $v_I^F(I;\theta_H) = a_2 + b_2 I$; $V_I(I;\theta_L) = c_1 + d_1 I$ and $v_I^F(I;\theta_L) = c_2 + d_2 I$.

Under FDI, since M_1 gains familiarity with FC with FC her investments at date 2 are given by (6a), which for the quadratic functional form reduces to:

$\frac{1}{2}(a_1 + b_1\hat{I}^F_{H1}) + \frac{1}{2}(a_2 + b_2\hat{I}^F_{H1}) = 1$, when $\theta = \theta_H$

This implies that:

(III) $\frac{1}{2}(a + b\hat{I}_{H1}) = 1$, where $a = a_1 + a_2$, and $b = b_1 + b_2$

Similarly, when $\theta = \theta_L$, we get:

(IV) $\frac{1}{2}(c + d\hat{I}^F_{L1}) = 1$, where $c = c_1 + c_2$, and $d = d_1 + d_2$

With exports, M_1 does not gain familiarity with the foreign country, and her investment levels are given by (7a). For the quadratic functional form, this implies that:

(V) $\frac{1}{2}\rho(a + b\tilde{I}_1^F) + \frac{1}{2}(1-\rho)(c + d\tilde{I}_1^F) = 1$

Now, multiplying (III) by ρ, and (IV) by $(1-\rho)$ and adding, we get:

(VI) $\frac{1}{2}\rho(a + b\hat{I}^F_{H1}) + \frac{1}{2}(1-\rho)(c + d\hat{I}^F_{L1}) = 1$

Equating (V) and (VI):

(VII) $\rho b\tilde{I}_1^F + (1-\rho)d\tilde{I}_1^F = \rho b\hat{I}^F_{H1} + (1-\rho)d\hat{I}^F_{H1}$

$$\Rightarrow \tilde{I}_1^F = \frac{\rho b\hat{I}^F_{H1} + (1-\rho)d\hat{I}^F_{L1}}{\rho b + (1-\rho)d}$$

This simplifies to:

(VIII) $\tilde{I}_1^F = \hat{I}^F_{L1} + \dfrac{\rho b}{\rho b + (1-\rho)d}[\hat{I}^F_{H1} - \hat{I}^F_{H1}]$

This implies that when $\rho = 0$ (or in other words when $\theta = \theta_L$), $\tilde{I}_1^F = \hat{I}_{L1}^F$ and when $\rho = 1$, $\tilde{I}_1^F = \hat{I}_{H1}^F$.

From this it follows that $\Pi_{FDI} = \Pi_E$ when $\rho = 0$ and when $\rho = 1$.

The question now is how does Π_{FDI} compare with Π_E for $\rho \in (0,1)$?

First of all we can show that the derivative of $\Pi_{FDI}(\rho)$ is independent of ρ. Differentiating (I) with respect to ρ, we get:

$$\text{(IX)} \qquad \frac{\partial \Pi_{FDI}(\rho)}{\partial \rho} = V(\hat{I}_{H1}^F;\theta_H) - V(\hat{I}_{L1}^F;\theta_L) - [\hat{I}_{H1}^F - \hat{I}_{L1}^F]$$

We have that $\Pi_{FDI} = \Pi_E$ when $\rho = 0$ and when $\rho = 1$. Since the slope of $\Pi_{FDI}(\rho)$ is a constant, we know that Π_E will have to exceed Π_{FDI} over some range of $\rho \in (0,1)$ *if* either:

$$\text{(i)} \qquad \left.\frac{\partial \Pi_E(\rho)}{\partial \rho}\right|_{\rho=0} > \left.\frac{\partial \Pi_{FDI}(\rho)}{\partial \rho}\right|_{\rho=0}$$

or

$$\text{(ii)} \qquad \left.\frac{\partial \Pi_E(\rho)}{\partial \rho}\right|_{\rho=1} < \left.\frac{\partial \Pi_{FDI}(\rho)}{\partial \rho}\right|_{\rho=1}$$

(i) and (ii) individually represent a sufficient condition for Π_E to exceed Π_{FDI}, and hence to counter Proposition 1.

We can now proceed to derive the conditions under which (i) and (ii) hold.

Differentiating (II) with respect to ρ, we get:

$$\frac{\partial \Pi_E(\rho)}{\partial \rho} = V(\tilde{I}_1^F;\theta_H) - V(\tilde{I}_1^F;\theta_L) + [\rho V_I(\tilde{I}_1^F;\theta_H) + (1-\rho)V_I(\tilde{I}_1^F;\theta_L)]\frac{\partial \tilde{I}_1^F}{\partial \rho} - \frac{\partial \tilde{I}_1^F}{\partial \rho}$$

To get an expression for $\frac{\partial \tilde{I}_1^F}{\partial \rho}$, we can differentiate (VIII), which was derived earlier from the FOCs, with respect to ρ:

$$\text{(X)} \qquad \frac{\partial \tilde{I}_1^F(\rho)}{\partial \rho} = \frac{bd[\hat{I}_{H1}^F - \hat{I}_{L1}^F]}{[\rho b + (1-\rho)d]^2}$$

Evaluating (X) at $\rho = 0$ gives:

$$\text{(XI)}\quad \left.\frac{\partial \tilde{I}_1^F(\rho)}{\partial \rho}\right|_{\rho=0} = \frac{b}{d}[\hat{I}_{H1}^F - \hat{I}_{L1}^F]$$

And evaluating at $\rho = 1$ yields:

$$\text{(XII)}\quad \left.\frac{\partial \tilde{I}_1^F(\rho)}{\partial \rho}\right|_{\rho=0} = \frac{d}{b}[\hat{I}_{H1}^F - \hat{I}_{L1}^F]$$

We therefore have that:

$$\text{(XIII)}\quad \left.\frac{\partial \Pi_E(\rho)}{\partial \rho}\right|_{\rho=0} = V(\hat{I}_{L1}^F;\theta_H) - V(\hat{I}_{L1}^F;\theta_L) + [V_I(\hat{I}_{L1}^F;\theta_L)]\frac{b}{d}[\hat{I}_{H1}^F - \hat{I}_{L1}^F] - \frac{b}{d}[\hat{I}_{H1}^F - \hat{I}_{L1}^F]$$

And:

$$\text{(XIV)}\quad \left.\frac{\partial \Pi_E(\rho)}{\partial \rho}\right|_{\rho=1} = V(\hat{I}_{H1}^F;\theta_H) - V(\hat{I}_{H1}^F;\theta_L) + [V_I(\hat{I}_{H1}^F;\theta_H)]\frac{d}{b}[\hat{I}_{H1}^F - \hat{I}_{L1}^F] - \frac{d}{b}[\hat{I}_{H1}^F - \hat{I}_{L1}^F]$$

From (IX) and (XIII) we get that for $\left.\frac{\partial \Pi_E(\rho)}{\partial \rho}\right|_{\rho=0} > \left.\frac{\partial \Pi_{FDI}(\rho)}{\partial \rho}\right|_{\rho=0}$ to hold:

$$\text{(i)}\quad \frac{b}{d}V_I(\hat{I}_{L1}^F;\theta_L) > \frac{V(\hat{I}_{H1}^F;\theta_H) - V(\hat{I}_{L1}^F;\theta_H)}{\hat{I}_{H1}^F - \hat{I}_{L1}^F} + (\frac{b}{d} - 1)$$

Similarly, from (IX) and (XIV) we can derive the expression for $\left.\frac{\partial \Pi_E(\rho)}{\partial \rho}\right|_{\rho=1} < \left.\frac{\partial \Pi_{FDI}(\rho)}{\partial \rho}\right|_{\rho=1}$:

$$\text{(ii)}\quad \frac{d}{b}V_I(\hat{I}_{H1}^F;\theta_H) < \frac{V(\hat{I}_{H1}^F;\theta_L) - V(\hat{I}_{L1}^F;\theta_L)}{\hat{I}_{H1}^F - \hat{I}_{L1}^F} + (\frac{d}{b} - 1)$$

So, given that transport costs and tariffs are zero, expressions (i) and (ii) are independent sufficient conditions for $\Pi_E > \Pi_{FDI}$. For positive, but sufficiently low values of transport costs and tariffs, it follows that exports can still yield a higher payoff for M_1 than FDI.

■

III. INCOMPLETE CONTRACTS AND THE FAILURE OF OIL FIELD UNITIZATION

1. Introduction

The characteristic feature of a majority of domestic oil reservoirs is the multiplicity of interests: numerous firms own leases that permit exploitation of a common pool of oil[1]. This, along with mobility of oil within the reservoir and the laws that determine ownership of the extracted oil, ensure that the oil industry is susceptible to common pool problems found in the competitive exploitation of many open-access natural resources. Numerous economists have argued, therefore, that the competitive extraction of oil is inefficient, and results in rent dissipation.

The existence of a common pool of oil in itself does not necessitate rent dissipation when oil is extracted competitively, and it is the interplay of multiple interests, the property rights over extracted oil, and the fluid nature of oil that guarantee rent dissipation. The law governing ownership of oil in the reservoir is the 'rule of capture': ownership of oil is granted to the entity that extracts it[2]. Oil is a mobile substance that flows from one part of the reservoir to another, and the degree of mobility is governed by the porosity of the surrounding rock formation and the viscosity of oil, which in itself depends to a large extent on the amount of natural gas dissolved in oil[3].

Given the migratory nature of oil and the rule of capture, there is an incentive for firms drilling competitively for oil to extract as much oil as possible before it is captured by a neighbor, leading to sub-optimally high rates of extraction. Rent dissipation then follows due to a number of reasons. As oil is rapidly removed from the reservoir, the natural gas trapped in the reservoir dissipates, and pressure within the reservoir drops. At the same time,

---

[1] Oil firms do not usually own the surface land above a reservoir. Rather, they obtain a lease from the surface owner that grants the firm the right to drill for oil in exchange for payments to the land-owner. These payments include an immediate cash payment, possible delay payments, and royalty payments on produced oil (McDonald (1971, pp. 13-14). It is the fragmentation of land ownership that leads to multiple oil firms possessing leases for oil extraction.

[2] Moreover, it is worth noting that the rule of capture applies to the flow of extracted oil, and not to the stock of oil in place. In other words, the first firm to extract oil from a reservoir does not capture the entire reservoir, only the oil extracted. If the rule of capture applied over the entire reservoir, the common pool problem would disappear. For a more detailed description of implications of first possession, see Lueck (1998).

[3] If oil were in immobile, solid substance, the ownership of surface land would dictate ownership of the subsurface mineral.

dissolved natural gas escapes from solution, resulting in an increase in the viscosity of the remaining oil. This necessitates the premature adoption of secondary recovery procedures, and leads to a loss of surplus. Further loss of surplus occurs as firms drill (usually redundant) wells along the boundaries of their leases to prevent oil from migrating to neighboring tracts.

Given the potential ill effects of competitive extraction, there is a rationale for oil firms to cooperate over the extraction of oil, that is, to find a common property arrangement that mitigates rent dissipation. The common property arrangement that is widely regarded as the solution to rent dissipation is unitization (McDonald (1971), Libecap and Wiggins (1985), Wiggins and Libecap (1985), Weaver (1986))[4]. Unitization is a contractual arrangement between the various participants in the development of the reservoir, and is, essentially, a joint venture between the oil-producing firms. A unitization contract typically involves two documents: the Unit Agreement between the oil producing firms (the lessees) and the owners of surface land (lessors), and the Unit Operating Agreement among the lessees themselves that outlines the key features regarding the operation of the unit (Libecap and Smith (1999), American Petroleum Institute (1993)).

While the agreement contains numerous details regarding the formation and operation of the unit, there are two important elements in a unitization agreement. The first is that all operation and production decisions are delegated to a designated unit operator; the remaining firms oversee the activities of the unit operator through a supervisory committee. The second is that the unitization contract specifies a sharing rule for substances produced and costs incurred by the unit. Unitization is, thus, a special type of joint venture where substances extracted from the reservoir and costs are divided among the firms on the basis of shares agreed upon in the unitization agreement. The shares specified in the initial contract between firms are once-and-for-all-shares and are seldom renegotiated. Moreover, the shares are over the output produced and costs incurred, and do not extend to the marketing of produced substances. Finally, the share of the output received by a firm equals its share of the total costs incurred by the firm.

[4] Complete unitization of the oil pool is not the only form of cooperation between firms in an oil field; other typical arrangements (some of which are enforced by state or federal governments) include prorationing agreements, well-spacing rules, partitioned development, and partial unitization agreements. These arrangements are usually viewed as being less effective solutions to the common pool problem compared to unitization.

Competitive extraction and unitization are, therefore, different organizational modes adopted by firms while extracting oil from a common pool. It is worthwhile to outline the differences between these organizational modes in terms of ownership and control of physical assets involved in the process of oil extraction. There are, in essence, three physical assets relevant for our analysis: a lease obtained by an oil firm from the land-owner, the wells used for extraction of oil, and the pool of oil itself. The ownership of a lease provides an oil firm with access to the reservoir, and hence the right to extract oil from the reservoir. With competitive extraction, any well built within the confines of the area prescribed by the lease is owned by the firm, and the firm maintains control rights over the use of the well. The ownership of the oil is dictated by the rule of capture; access to the reservoir does not guarantee a firm ownership of oil. With unitization, each firm retains ownership of the lease, but transfers ownership of its wells to the unit[5]. The firm thus *transfers control rights* over the production process to the unit and more specifically, to the unit operator. Finally, ownership of oil is determined on the basis of shares agreed upon in the *ex ante* unitization contract.

The crucial difference, then, between the two organizational modes is the transfer of ownership of wells with a unitization agreement, and differing ownership rules over the extracted oil. As Grossman and Hart (1986), Hart and Moore (1990) and Hart (1995) suggest, ownership is important when contracts are incomplete, as such a contract confers the owner of an asset residual control rights over the use of the asset. The sub-terranean nature of oil deposits provides a natural setting for contractual incompleteness. Numerous features of the reservoir, such as the nature of the rock formation, the exact dimensions of the reservoir, the quantity of recoverable oil, the strength of the drive mechanism, the dynamics of oil migration within the reservoir, etc., remain unknown till the actual act of extraction. It is virtually impossible, therefore, for firms to contract upon a specific production technology or a specific quantity of oil to be extracted in an *ex ante* contract. The unit operator is left

[5] This supposes that unitization occurs after the construction of wells. The wells are taken over by the unit in return for payments for any expenditure the firm may have incurred in constructing the well. Other investments in infrastructure, such as pipelines built by a firm prior to unitization, are compensated for in a similar manner. When wells or pipelines are constructed after the unitization agreement, the costs are defrayed according to the sharing rule in the agreement (see Libecap and Smith (1999), for example).

with considerable residual control rights while conducting day-to-day operations[6].

The rent dissipation with competitive extraction can be considerable and run into billions of dollars of lost revenues (Libecap and Wiggins (1985)). Given the magnitude of the potential gains due to unitization, a puzzle that exists in the oil industry, and one that this paper attempts to address, is the reluctance of firms to voluntarily enter a unitization agreement. The study by Libecap and Wiggins (1985) shows that in 1975, only 38 percent of oil production in Oklahoma and 20 percent of oil production in Texas were accounted for by unitized fields. This has often induced the federal and state government to step in and pass statutory unitization laws; the presumption behind such regulatory policies being that unitization is a more efficient method of extracting oil.

The literature attempting to explain the puzzle usually presupposes that unitization is efficient, and looks for frictions in the real world to explain why an inefficient organizational mode (such as competitive extraction) may be chosen. McDonald (1971) and Weaver (1986) suggest a number of reasons: ignorance, mistrust, difficulty in contacting owners, fear of reduced current income, fear of violating antitrust laws, and fear of increased legal problems. While any or all of these factors may be true, they are certainly not insurmountable; misplaced fears and mistrust are not particularly difficult to dispel when large gains in profits are to be made. More viable economic rationales suggested by the two authors include pride of ownership and operational control, loss of operating experience, and holdouts due to structural advantage. Oil and natural gas in the reservoir are migratory and this may result in oil and gas being driven to some wells as drilling progresses, away from others; such wells are said to have a structural advantage.

Wiggins and Libecap (1985) argue that unitization is observed infrequently as a result of breakdown in bargaining during unitization contract negotiations. The unitization agreement incorporates sharing rules over division of the oil produced from the reservoir and costs incurred by the unit while producing the oil. Given that each firm has more information than other firms about the amount of oil it is capable of extracting from its tract,

[6] As Libecap and Smith (1999) state: "Certain contingencies are laid out...but day to day operation of the unit is left unspecified in the contract, with the unit operator bound only to a "best efforts" standard of performance...It is not possible ex ante to define strict performance criteria for the unit operator because far too much remains unknown at the time the contract is written..."

and hence the value of its lease, firms may fail to reach an agreement regarding the sharing rule. In other words, the breakdown in bargaining over the terms in the unitization contract occurs due to asymmetric information, as firms find it optimal to delay entering the unitization agreement till the asymmetry is narrowed over time with the arrival of new information. Other notable contributions include Lueck and Schenewerk (1996), which argues that the transaction costs involved with a large number of contracting parties may erode gains from unitization, and Smith (1987), which highlights the role of small interests in delaying unit formation.

Libecap and Smith (2001) continue to assume that unitization leads to a higher expected surplus than other modes of exploiting the reservoir, and analyze the impact of the reservoir having two assets (oil and natural gas). Specifically, they argue that unitization involves firms trading their initial endowment of the two assets, where their share of oil may be different from their share of natural gas, for fixed and common shares of oil and natural gas produced from the entire unit. If, as is the case in reality, the two assets have differing values and are subject to different levels of uncertainty, risk attitudes (or, alternatively, differing expectations regarding the value of the assets) may imply that unitization is not Pareto improving relative to certain other modes for exploiting oil[7], and some firms may be unwilling to make the trade implicit in unitization agreements.

This paper questions the assumption that unitization is always surplus enhancing, and argues that the incompleteness of unitization contracts may result in competitive extraction resulting in a higher surplus. Two factors play a prominent role in the developing the importance of contractual incompleteness in this paper. First is the fact that operations within a particular reservoir can have external, typically non-contractible, benefits that are realized as a by-product of extracting oil from a reservoir. Examples of such benefits include the value of operating experience that firms may find useful when exploiting their leases in other reservoirs; the ability to strengthen existing relationships with downstream refineries, and arrive at preferential trading arrangements; the ability to realize economies of scale and compete with other firms, etc. The unitization agreement prescribes shares over the value of the output extracted from the reservoir and these non-contractible benefits are not subject to

[7] Libecap and Smith focus on the choice between partitioned development of the oil field and unitization.

the same sharing rule. The second factor is the transfer of control rights to the unit operator, and the inability to contract on the efforts of the unit operator in developing the reservoir.

In contrast to Wiggins and Libecap (1985), who stress the role of *ex ante* bargaining in precipitating a breakdown of negations, this paper focuses on the role of *ex post* bargaining. The incompleteness of the unitization contract, arising due to the inability to prescribe state contingent activities of the unit operator and to devise *ex ante* sharing rules over non-contractible benefits, implies that firms will renegotiate over the division of the total surplus realized from unitization. So, firms other than the unit operator will capture a part of the gains resulting from increased effort from the unit-operator. This acts as a disincentive for the unit-operator to exert effort. Moreover, there will typically be an inefficiency associated with unitization regarding the quantity of oil extracted from the reservoir.

The paper identifies two factors relevant for the choice between competitive extraction and unitization. The first is the common pool problem, which results in competitive extraction being inefficient. The second is the presence of non-contractible benefits and the incomplete nature of unitization contracts, which imposes inefficiencies on unitization. We thus show that both competitive extraction and unitization involve inefficiencies, and in certain instances, the inefficiency with the former may be lower than that associated with the latter; consequently, competitive extraction may result in a higher surplus. While there is enough evidence to support the fact that the *ex ante* bargaining problem is an important factor in the failure of unitization agreements, this paper suggests that the anticipation of *ex post* bargaining can play an important role its failure as well, and can aggravate the unattractiveness of participating in unitization agreements. The presumption that unitization results in a higher surplus than competitive extraction is not innocuous when unitization contracts are incomplete.

The paper is organized as follows. The general model is described in Section 2, and the equilibrium outcomes under unitization and competitive extraction are solved for in Section 3. Section 4 considers a specific version of the general model, and highlights the existence of inefficiencies with both organizational modes. Section 5 presents extensions, and concludes.

2. The model framework

Two firms, F_1 and F_2, operated by risk-neutral managers M_1 and M_2[8], respectively, own leases that provide access to a reservoir, Z. The reservoir is characterized by a common pool of oil shared between the two firms. Ownership of a lease provides a firm with the right to explore and drill for oil by constructing oil wells within the area described by the lease. In what follows, we shall assume that drilling wells are already in place, and that each firm has built a single well on its land. There exists, therefore, a one to one correspondence between the number of leases and the number of wells, allowing for a convenient notation where F_1 (F_2) owns lease A (B), on which a drilling well A (B)[9] has been constructed. F_1 and F_2 can organize extraction of oil from the common pool in different ways[10]:

(a) *Competitive extraction*: Each firm drills oil competitively, without coordinating its activities with the other firm. This would imply that F_1 (F_2) drills using well A (B), and appropriates all the oil produced from well A (B).

(b) *Unitization*: Firms can cooperate over the extraction of oil by signing a unitization agreement. The typical unitization agreement selects a unit operator who is responsible for operating both wells, and assigns a share to each firm (α to F_1 and $(1-\alpha)$ to F_2) over contractible variables. Unitization agreements observed in the real world split oil recovered from the field and the costs incurred during extraction on the basis of the prescribed shares. We shall assume throughout the paper, therefore, that quantities of oil and operating costs are (*ex post*) contractible. We also assume the following:

Assumption 1:

(i) The shares prescribed in an initial unitization agreement are not subject to renegotiation later on.

(ii) The share of oil a firm receives is equal to the share of operating costs that it pays.

---

[8] Since we shall not be concerned with integration among firms, the notation for a firm and its corresponding manager shall be used interchangeably throughout.

[9] Throughout this paper, subscripts '*1*' and '*2*' shall be used to denote firms, and subscripts '*A*' and '*B*' to denote wells. Moreover, we shall employ the subscript '*m*' to describe an arbitrary well, and '*n*' to denote the other well; similarly an arbitrary firm or manager is identified with the subscript '*k*', with '*l*' referring to the other manager.

[10] This list is by no means exhaustive; other organizational modes for extracting oil such as prorationing agreements, partitioned development and partial unitization are not considered here.

(iii) Firms are not concerned about issues other than surplus maximization while choosing shares in a unitization agreement.

(iv) Firms cannot opt out of the unitization agreement once it is in place[11].

We consider a three period model with the following features. At $t = 0$, the firms have the option of entering into a unitization contract that has the characteristics described earlier. The contract may also prescribe lump-sum transfers from one firm to another[12]. The default situation where the firms do not explicitly agree to cooperate, results in competitive extraction.

At $t = 1$ (the *ex ante* period), the managers undertake efforts (or investments) that enhance the value of the reservoir operations. These actions are non-contractible[13], which is a natural assumption to make in the context of oil extraction, as outlined in section 1[14]. The efforts of M_1 and M_2 are denoted by i_1 and i_2, respectively, while the efforts specific to wells A and B are denoted by i_A and i_B. How the efforts of the managers are distributed among the wells depends on the organizational mode chosen at $t = 0$. Under competitive extraction, each firm has access only to the well on its lease, and therefore $i_1 = i_A$ and $i_2 = i_B$. With a unitization agreement, operating rights of both wells are handed over to the unit operator, and it is reasonable to expect that the unit operator is responsible for the important efforts on both wells. In this case, if M_k is appointed the unit operator, $i_k = i_A + i_B$, and $i_l = 0$ [15]. This is a departure from the standard analysis (Grossman and Hart (1986), Hart and Moore (1990), Hart (1996), etc.), where it is assumed that even after a change in the

[11] The unitization agreement may, for example, prescribe a large penalty for a firm that decides to opt out. This implicitly requires that courts can observe a firms decision to leave the unit.

[12] Unitization contracts often involve side payments, usually to compensate a firm for past investment outlays.

[13] These efforts are assumed to be observable, but non-verifiable.

[14] The efforts can represent, for example, investments in gathering information about the structure of the reservoir in the neighborhood of a well, efforts in devising appropriate recovery procedures, investments in understanding the common pool problem, and so on.

[15] The fact that the effort of the non-unit operator is normalized to zero is a simplifying assumption. It is conceivable that M_l may make some effort at t=1 even if M_k is the unit operator. Ultimately, however, only an operator of a well has the ability, by virtue of being responsible for the day-to-day operations of the well, to make important value enhancing investments. We note in passing here that this leads to a bias against unitization since we are substituting the two managers who function under competitive extraction with one manager under unitization. This bias can be artificially inflated to (theoretically) force competitive extraction to result in a higher surplus. This bias is eliminated in Section 4.

organizational structure, both managers continue to make valuable contributions to the productivity of the assets. While this is a reasonable assumption in the context of integration[16], which is the focus of these papers, it is clearly unrealistic in our current context, where operating rights are contractually transferred to the unit-operator in the unitization agreement. The non-contractible costs of exerting effort for the two managers are denoted $C_1(i_1)$ and $C_2(i_2)$, respectively.

Various features of the reservoir, such as the nature of the underlying rock formation, the quantity and value of recoverable oil recoverable oil from its subterranean location, the strength of the drive mechanism, the dynamics of oil migration within the reservoir, etc., are unknown *ex ante* when the investments are made[17]. This is resolved, and the state of the world, is determined between $t = 1$ and $t = 2$.

At $t = 2$ (the *ex post* period), oil is extracted from the two wells, possibly after some costless renegotiation between the two firms, and benefits are realized. The timeline is summarized below.

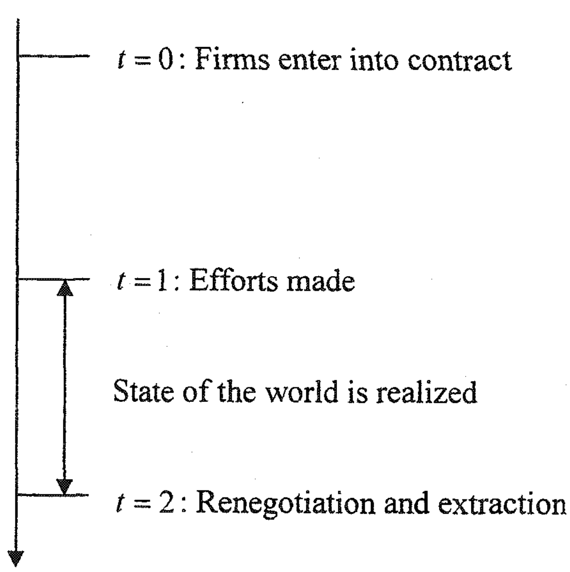

Figure 3.1: The timeline

[16] The manager of the firm that is acquired is assumed to remain in the (integrated) organization as an employee.

[17] The reader will recall from Section 1 that these are the features that motivated the incompleteness of the unitization contract.

The quantity of oil extracted from well $m \in \{A, B\}$ is denoted q_m. We assume that quantities of oil are *ex post* contractible (at $t = 2$), but are non-contractible at $t = 0$ due to the uncertain nature of the reservoir characteristics at that time. The underlying idea is that it is prohibitively expensive to specify quantities of oil as a function of the all possible realizations of the state of the world[18]. Once the relevant state is known, however, firms will negotiate over quantities of oil extracted from each well at $t = 2$ for the given state *if* there exists a contract at $t = 0$ allowing for coordination of the activities at a later stage. If a unitization agreement is in place, the quantity of oil extracted is negotiated, and then distributed among the firms as specified by the initial contract (that is, shares $(\alpha, 1-\alpha)$). With competitive extraction, there is no cooperation between the firms at $t = 2$[19]. We assume that there are two types of benefits realized at $t = 2$ from operating reservoir Z:

(a) *Contractible benefits:* This refers to the direct value of the oil recovered from reservoir; in other words, it refers to profits. The contractible benefit from operating well m is $V_m(q_m, q_n, i_m)$, where $m, n \in \{A, B\}$. Later on we will find it useful to assume a specific form of the contractible benefits; that form is introduced at this stage to add to the intuition:

(1) $$V_m(q_m, q_n, i_m) = pq_m - C_m(q_m, q_n, i_m), \forall m, n \in \{A, B\}, m \neq n$$

Here, p denotes the (assumed fixed) price at which the quantity of oil obtained from well m can be sold, and C_m the cost of operating the well[20]. As with the quantity of oil, we shall assume that C_m is observable and verifiable at $t = 2$, but is non-contractible at $t = 0$. When there is a unitization agreement in place, *total costs* are divided between the firms according to the shares $(\alpha, 1-\alpha)$. The formulation in (1) implies that the efforts at $t = 1$ affect the operating costs at $t = 2$.

[18] There are numerous variables that determine the quantity of oil that can be recovered from a reservoir, and the sub-terannean location ensures that these variables cannot be determined before the reservoir is exploited and considerable effort has gone into understanding reservoir characteristics (in our model, these efforts are made at $t = 1$). A contract can, of course, prescribe an arbitrary quantity of oil to be extracted, but this is bound to be renegotiated as more information about the nature of the reservoir comes in. In other words, firms cannot contractually commit to extracting a particular quantity of oil.

[19] Competitive extraction is *defined* as the absence of cooperation throughout the life of the reservoir. There is no inherent reason why firms cannot proceed to cooperate at $t = 2$ in spite of the absence of a contract at $t = 0$. This, however, does not conform to the nature of competitive extraction.

[20] Note that i_A and i_B are *well specific*, and *not* manager specific, that is, it does not matter which manager puts in the effort into a particular well. The realized cost reduction due to a particular level of investment would therefore be the same for either manager operating on a specific well.

Assumption 2: $\forall m,n \in \{A,B\}, m \neq n$,

$C_m(q_m, q_n, i_m)$ is (strictly) convex, (strictly) increasing in q_m and (strictly) decreasing in i_m. This implies that $V_m(q_m, q_n, i_m)$ is (strictly) concave and (strictly) increasing in i_m.

Assumption 3 (The common pool problem): $\forall m,n \in \{A,B\}, m \neq n$,

(i) $C_m(q_m, q_n, i_m)$ is (strictly) increasing in q_n, that is, $\dfrac{\partial V_m(q_m, q_n, i_m)}{\partial q_n} < 0$

(ii) $\dfrac{\partial^2 V_m(q_m, q_n, i_m)}{\partial q_m \partial q_n} < 0$

It is evident that with competitive extraction, the contractible benefit accruing to F_1 is $V_A(q_A, q_B, i_A)$ and that to F_2 is $V_B(q_A, q_B, i_B)$, with $i_1 = i_A$ and $i_2 = i_B$. On the other hand, with a unitization contract in place, F_1 receives a portion of the contractible benefit equal to $\alpha[V_A(q_A, q_B, i_A) + V_B(q_A, q_B, i_B)]$, with F_2 receiving the remainder.

(b) *Non-contractible benefits*: There may exist other, non-contractible, benefits that accrue to the firms. These arise from the effect that conducting operations on one particular reservoir have on various *other* activities of the firms. These external benefits cannot, typically, be shared in the same manner as the surplus from the reservoir in focus[21]. Let the non-contractible benefit received by F_k be denoted $H_k(q_k, i_k), \forall k \in \{1,2\}$, which is strictly concave, and increasing in both its arguments. The following assumption is not crucial to the subsequent analysis, and helps to simplify matters:

Assumption 4: $, \forall k \in \{1,2\}$

$H_k(q_k, i_k) = R_k(q_k) + B_k(i_k)$, that is, it is additively separable.

The oil that M_k receives from Z has non-contractible benefits, and this is captured by $R_k(q_k)$. For example, the ability of F_k to compete with other firms in downstream and

[21] In a unitization agreement, firms have to determine the division of the non-contractible surplus during *ex post* renegotiation. With competitive extraction, firms do not negotiate the division of these benefits.

lateral markets may depend on the total quantity of oil handled by the firm from all its leases[22]. At the same time, M_k may derive non-contractible benefits from exerting effort, $B_k(i_k)$, such as the operating experience from drilling and operating wells on the focus reservoir, which allows M_k to operate her other leases more effectively.

In this paper, we shall also assume that all bargaining power lies with one firm at $t = 0$; so one firm offers a take-it-or-leave-it contract to the other firm, which accepts/rejects the offer. There is no *ex ante* bargaining over unit shares. The firm offering the contract will maximize the total surplus, and offer the other firm its reservation price (which could be its surplus under competitive extraction).

We do, however, allow for bargaining during negotiations at $t = 2$, which follows the protocol of Rubinstein (1982). Suppose that M_1 is appointed as the unit operator in the initial contract. When agreeing to appoint M_1 as the sole operator of the reservoir M_2, essentially, hands over residual control rights over drilling decisions to M_1. As pointed out in Grossman and Hart (1986), residual control rights matter when contracts are incomplete. In our context, if firms were to disagree during negotiations at $t = 2$, the default situation is one where M_1 non-cooperatively chooses quantities at $t = 2$ to maximize her private benefits. The specification of $(\alpha, 1-\alpha)$ in the initial contract allows firms to split the resultant oil and costs. The firms therefore receive flows of income while bargaining continues; these flows are thus inside options for the firms.

Given that the quantity of oil is *ex post* contractible, however, firms will negotiate the extraction levels of oil. The assumption that firms cannot opt out of the unitization agreement at $t = 2$ plays an important role here and restricts the *ex post* alternatives available to the firms[23]. The outcome of the bargaining process corresponds to Nash's split the difference rule, where each firm receives its inside option plus one half of the gains from trade.

[22] $R_k(q_k)$ can also reflect any private benefit the managers get from empire building, being recognized as an oil-tycoon, etc. Alternatively, $R_k(q_k)$ may indicate the value from handling the oil produced from a given field for marketing purposes; a firm may be able to establish relationships with downstream firms while operating the focus reservoir that are of value to the firm in operations that has in other fields.

[23] More specifically, it eliminates outside options available to firms.

3. Efficiency, competitive extraction and unitization

3.1 The efficient outcome

We can begin by solving for the first best levels of surplus, which occurs when there are no non-contractible variables. The solution is obtained by backward induction. Let $Q = q_A + q_B$ represent the total quantity of oil extracted from the reservoir. At $t = 2$, the firms attempt to co-operatively maximize the *ex post* surplus[24]:

(2) $$\max_{q_A, q_B, \gamma} V_A(q_A, q_B, i_A) + V_B(q_A, q_B, i_B) + R_1(\gamma Q) + R_2((1-\gamma)Q)$$

Here, $q_1 = \gamma Q$ is the amount of oil distributed to F_1, while $q_2 = q_A + q_B - q_1 = (1-\gamma)Q$ is the amount of oil given to F_2. V_A and V_B are defined according to equation (1):

$$V_m(q_m, q_n, i_m) = pq_m - C_m(q_m, q_n, i_m), \forall m, n \in \{A, B\}, m \neq n$$

The necessary and sufficient conditions for an interior solution are[25]:

(2a) $$\frac{\partial V_m}{\partial q_m} + \frac{\partial V_n}{\partial q_m} + \gamma \frac{\partial R_1}{\partial q_1} + (1-\gamma)\frac{\partial R_2}{\partial q_2} = 0, \ \forall m, n \in \{A, B\}, m \neq n$$

(2b) $$\frac{\partial R_1}{\partial q_1} - \frac{\partial R_2}{\partial q_2} = 0$$

Let the solution to (2a) and (2b) be (q_A^*, q_B^*, γ^*). It is evident from (2a) and (2b) that at $t = 2$, the efficient outcome satisfies:

(2c) $$\frac{\partial V_A}{\partial q_A} + \frac{\partial V_B}{\partial q_A} = \frac{\partial V_A}{\partial q_B} + \frac{\partial V_B}{\partial q_B} = -\frac{\partial R_1}{\partial q_1} = -\frac{\partial R_2}{\partial q_2}$$

This implies that the efficient outcome requires that at $t = 2$, the wells A and B are utilized till their impact on marginal contractible benefits from the reservoir are equalized, that oil is divided between F_1 and F_2 such that the marginal non-contractible benefit to each firm is equalized, and that the sum of marginal contractible and non-contractible benefits from operating extracting more oil from a well equals its marginal cost. The efficient *ex post* surplus for given levels of investments made at $t = 1$ is:

(2d) $$S^* = V_A(q_A^*, q_B^*, i_A) + V_B(q_A^*, q_B^*, i_B) + R_1(q_1^*) + R_2(q_2^*) + B_1(i_1) + B_2(i_2)$$

---

[24] Note that $B_k(i_k)$ is pre-determined at $t = 2$, and therefore is not explicitly addressed in the problem.

[25] For notational brevity, the arguments of the various functions are suppressed except where necessary.

At $t=1$, F_1 and F_2 have to decide on the optimal effort to exert on each well, and the optimal distribution of effort amongst the two managers. That is, they solve:

(3) $$\max_{i_A, i_B, i_1} S^* - C_1(i_1) - C_2(i_2)$$

Recall that $q_2^* = q_A^* + q_B^* - q_1^*$, and $i_2 = i_A + i_B - i_1$. Using the envelope theorem and assuming an interior solution, the first order conditions are:

(3a) $$\frac{\partial V_m}{\partial i_m} + \frac{\partial B_2}{\partial i_2} - \frac{\partial C_2}{\partial i_2} = 0,\ \forall m \in \{A, B\}$$

(3b) $$\frac{\partial B_1}{\partial i_1} - \frac{\partial C_1}{\partial i_1} - \frac{\partial B_2}{\partial i_2} + \frac{\partial C_2}{\partial i_2} = 0$$

Together, these equations imply that:

(3-c) $$\frac{\partial V_A}{\partial i_A} = \frac{\partial V_B}{\partial i_B} = \frac{\partial C_2}{\partial i_2} - \frac{\partial B_2}{\partial i_2} = \frac{\partial C_1}{\partial i_1} - \frac{\partial B_1}{\partial i_1}$$

Let the solution to (3c) be (i_A^*, i_B^*, i_1^*). So, efficiency requires that at $t=1$ efforts are allocated among wells and firms in a manner such that the marginal contractible benefit from allocating more effort to a well is equalized across wells, that effort is allocated between firms so as to equalize (net) marginal non-contractible benefits, and that the sum of marginal contractible and non-contractible benefits from exerting effort equals the marginal cost to a firm of exerting effort. The expected surplus at $t=0$ is given by:

(4) $$\begin{aligned}\Pi^* &= V_A(q_A^*, q_B^*, i_A^*) + V_B(q_A^*, q_B^*, i_B^*) + R_1(q_1^*) + R_2(q_2^*) \\ &\quad + B_1(i_1^*) + B_2(i_2^*) - C_1(i_1^*) - C_2(i_2^*)\end{aligned}$$

3.2 Competitive extraction

We now assume that investments cannot be contracted upon. Recall that competitive extraction is characterized by a lack of cooperation, implying that $i_1 = i_A$, $i_2 = i_B$, $q_1 = q_A$ and $q_2 = q_B$.

At $t=2$, the objectives of M_1 and M_2 are described by equations (5a) and (5b), respectively.

(5a) $$\max_{q_A} V_A(q_A, q_B, i_A) + R_1(q_A)$$

(5b) $$\max_{q_B} V_B(q_A, q_B, i_B) + R_2(q_B)$$

This yields the following equilibrium conditions at $t = 3$ for given levels of investments:

(5c) $$\frac{\partial V_A}{\partial q_A} + \frac{\partial R_1}{\partial q_1} = 0$$

(5d) $$\frac{\partial V_B}{\partial q_B} + \frac{\partial R_2}{\partial q_2} = 0$$

Let the equilibrium quantities that solve (5c) and (5d) be (q_A^c, q_B^c); the superscript 'c' indicates the competitive outcome. The payoffs received by the firms, and the total surplus at $t = 2$, are:

(6a) $$S_1^c = V_A(q_A^c, q_B^c, i_A) + R_1(q_A^c) + B_1(i_A)$$

(6b) $$S_2^c = V_B(q_A^c, q_B^c, i_B) + R_2(q_A^c) + B_2(i_B)$$

(6c) $$S^c = S_1^c + S_2^c$$
$$= V_A(q_A^c, q_B^c, i_A) + V_B(q_A^c, q_B^c, i_B) + R_1(q_A^c) + R_2(q_B^c) + B_1(i_A) + B_2(i_B)$$

Proceeding backwards, we can now derive the optimal effort levels at $t = 1$. This involves M_1 choosing effort to maximize (7a), and M_2 maximizing (7b).

(7a) $$\max_{i_A} V_A(q_A^c, q_B^c, i_A) + R_1(q_A^c) + B_1(i_A) - C_1(i_A)$$

(7b) $$\max_{i_B} V_B(q_A^c, q_B^c, i_B) + R_2(q_B^c) + B_2(i_B) - C_2(i_B)$$

Taking the derivatives, and using (5c) and (5d), we get:

(7c) $$\frac{\partial V_A}{\partial q_B}\frac{\partial q_B^c}{\partial i_A} + \frac{\partial V_A}{\partial i_A} + \frac{\partial B_1}{\partial i_1} - \frac{\partial C_1}{\partial i_1} = 0$$

(7d) $$\frac{\partial V_B}{\partial q_A}\frac{\partial q_A^c}{\partial i_B} + \frac{\partial V_B}{\partial i_B} + \frac{\partial B_2}{\partial i_2} - \frac{\partial C_2}{\partial i_2} = 0$$

Let the solution to (7c) and (7d) be denoted (i_A^c, i_B^c). Competitive extraction therefore yields an expected surplus at $t = 0$ of:

(7e) $$\Pi^c = V_A(q_A^c, q_B^c, i_A^c) + V_B(q_A^c, q_B^c, i_B^c) + R_1(q_A^c) + R_2(q_B^c) + B_1(i_1^c) + B_2(i_2^c) - C_1(i_1^c) - C_2(i_2^c)$$

Of the total surplus, F_1 obtains:

(7f) $\Pi_1^c = V_A(q_A^c, q_B^c, i_A^c) + R_1(q_A^c) + B_1(i_1^c) - C_1(i_1^c)$

Similarly, F_2 receives an *ex ante* payoff of:

(7g) $\Pi_2^c = V_B(q_A^c, q_B^c, i_B^c) + R_2(q_B^c) + B_2(i_2^c) - C_2(i_2^c)$

Comparing results from competitive extraction with that of the efficient outcome, if the common pool problem disappears (that is, the conditions $\frac{\partial V_m(q_m, q_n, i_m)}{\partial q_n} = 0$ and $\frac{\partial^2 V_m(q_m, q_n, i_m)}{\partial q_m \partial q_n} = 0$ hold instead of Assumption 3), competitive extraction will satisfy the condition for efficient quantities being extracted from each well at $t = 2$ (see (2a), (2b), (5a) and (5b)), and efficient investments being made on each well $t = 1$ (see (3a), (3b), (7c) and (7d)). However, the distribution of oil and the allocation of efforts among firms under competitive extraction may not be efficient due to the constraint that $i_1 = i_A$, $i_2 = i_B$, $q_1 = q_A$ and $q_2 = q_B$.

3.3 Unitization

Unlike competitive extraction, where there is no cooperation between the firms throughout the extractive process, a unitization agreement at $t = 0$ allows firm to negotiate the quantity of oil drilled from the two wells at $t = 3$. While negotiating quantity levels, firms also have to decide how to divide the *ex post* non-contractible benefits; the contractible benefits and costs are split according to shares prescribed in the initial unitization agreement.

We have assumed that bargaining follows Rubinstein's alternating-offers procedure; the bargaining outcome corresponds to the Nash bargaining solution with each firm receiving its inside option plus one half of the gains from trade.

Without loss of generality, suppose that M_1 is appointed the unit operator at $t = 0$[26]. Consider first the payoffs to each firm at $t = 3$ if bargaining fails. M_1, by virtue of being the appointed the unit operator, has residual control rights over the quantities of oil to be

[26] The case where M_2 is the unit operator is analyzed analogously.

extracted from the wells, and will therefore solve, for given levels of investments:

(8) $$\max_{q_A,q_B} \alpha[V_A(q_A,q_B,i_A)+V_B(q_A,q_B,i_B]+R_1(\alpha(q_A+q_B))$$

This results in the first order conditions:

(8a) $$\frac{\partial V_A}{\partial q_m}+\frac{\partial V_B}{\partial q_m}+\frac{\partial R_1}{\partial q_1}=0, \forall m \in \{A,B\}$$

Let the solution to the equations implied by (8-a) be (q_A^d, q_B^d), where the superscript '*d*' implies disagreement between the two firms. Letting $Q^d = q_A^d + q_B^d$, the payoff received by the two firms, and the total *ex post* surplus in this case are given by the following[27]:

(9a) $$S_1^d = \alpha[V_A(q_A^d,q_B^d,i_A)+V_B(q_A^d,q_B^d,i_B)]+R_1(\alpha(Q^d))+B_1(i_1)$$

(9b) $$S_2^d = (1-\alpha)[V_A(q_A^d,q_B^d,i_A)+V_B(q_A^d,q_B^d,i_B)]+R_2((1-\alpha)(Q^d))$$

(9c) $$S^d = S_1^d + S_2^d$$

$$= V_A(q_A^d,q_B^d,i_A)+V_B(q_A^d,q_B^d,i_B)+R_1(\alpha Q^d)+R_2((1-\alpha)Q^d)+B_1(i_1)$$

Clearly (8a) does not, in general, maximize the ex post surplus, so firms can gain by negotiating different quantities extracted for any given state. That is, they can solve[28]:

(10) $$\max_{q_A,q_B} V_A(q_A,q_B,i_A)+V_B(q_A,q_B,i_B)+R_1(\alpha(q_A+q_B))+R_2((1-\alpha)(q_A+q_B))$$

The first order conditions are:

(10a) $$\frac{\partial V_A}{\partial q_m}+\frac{\partial V_B}{\partial q_m}+\alpha\frac{\partial R_1}{\partial q_1}+(1-\alpha)\frac{\partial R_2}{\partial q_2}=0, \ \forall m \in \{A,B\}$$

Let the solution be denoted (q_A^{u1}, q_B^{u1}), where the superscript '*u1*' denotes agreement between the two firms with F_1 selected as the unit operator. Letting $Q^{u1} = q_A^{u1} + q_B^{u1}$, the resultant surplus at $t = 3$ for any given state is:

(11) $$S^{u1} = V_A(q_A^{u1},q_B^{u1},i_A)+V_B(q_A^{u1},q_B^{u1},i_B)+R_1(\alpha Q^{u1})+R_2((1-\alpha)(Q^{u1})+B_1(i_1)$$

Since $S^{u1} > S^d$, renegotiation leads to gains equal to $S^{u1} - S^d$. The payoff to each firm,

---

[27] Since $i_2 = 0$, $B_2(i_2) = 0$.

[28] Note that (10) differs from (2); firms are now constrained by a contract that imposes a specific distribution of the oil, rather than their being able to decide *ex post* the efficient distribution for a given realization of the random variable. In general, therefore, the solution to (10) will fail to result in efficient quantities of oil being extracted.

following the Nash bargaining solution is:

(11a) $S_1^{u1} = S_1^d + \frac{1}{2}[S^{u1} - S^d] = \frac{1}{2}[S^{u1} + S_1^d - S_2^d]$

(11b) $S_2^{u1} = S_2^d + \frac{1}{2}[S^{u1} - S^d] = \frac{1}{2}[S^{u1} + S_2^d - S_1^d]$

At $t = 1$, the unit operator makes efforts (non-cooperatively) so as to maximize the expected value of her *ex post* payoff. The problem facing M_1 at $t = 1$ is therefore:

(12) $\max_{i_A, i_B} S_1^{u1} - C_1(i_1)$, where $i_1 = i_A + i_B$

Taking the partial derivatives, and using (8a) and (10a) to simplify, the first order conditions for the maximization problem (12) reduce to, $\forall m \in \{A, B\}$:

(12a) $$\frac{1}{2}\{\frac{\partial V_m^{u1}}{\partial i_m} + (2\alpha - 1)\frac{\partial V_m^d}{\partial i_m} + (1-\alpha)[\frac{\partial R_1^d}{\partial q_1} - \frac{\partial R_2^d}{\partial q_2}][\frac{\partial q_A^d}{\partial i_m} + \frac{\partial q_B^d}{\partial i_m}]\} + \frac{\partial B_1}{\partial i_1} - \frac{\partial C_1}{\partial i_1} = 0$$

Here, the simplifying notation adopted is: $V_m^{u1} = V_m(q_A^{u1}, q_B^{u1}, i_m)$, $V_m^d = V_m(q_A^d, q_B^d, i_m)$, $R_1^d = R_1(\alpha Q^d)$ and $R_2^d = R_2((1-\alpha)Q^d)$.

Let the effort levels implied by (12-a) be (i_A^{u1}, i_B^{u1}).

At $t = 1$, for a given $(\alpha, 1-\alpha)$ and M_1 being chosen as the unit operator, the expected surplus is:

(13) $$\begin{aligned}\Pi^{u1}(\alpha) &= V_A(q_A^{u1}, q_B^{u1}, i_A^{u1}) + V_B(q_A^{u1}, q_B^{u1}, i_B^{u1}) + R_1(\alpha Q^{u1}) \\ &+ R_2((1-\alpha)Q^{u1}) + B_1(i_1^{u1}) - C_1(i_1^{u1})\end{aligned}$$

At $t = 0$, there are two decisions that have to be made: the allocation of shares to the firms, and the choice of the unit operator. Note that we have not placed any restrictions on the allocation of shares, and simply require that they maximize the *ex ante* surplus[29]. The shares $(\alpha, 1-\alpha)$ are chosen to maximize Π^{u1}. Given that M_1 is the unit operator, the choice of α will maximize equation (13), and would yield the following first order condition for an interior solution:

(14) $$(\frac{\partial V_A}{\partial i_A} + \frac{\partial B_1}{\partial i_1} - \frac{\partial C_1}{\partial i_1})\frac{\partial i_A^{u1}}{\partial \alpha} + (\frac{\partial V_B}{\partial i_B} + \frac{\partial B_1}{\partial i_1} - \frac{\partial C_1}{\partial i_1})\frac{\partial i_B^{u1}}{\partial \alpha} + Q^{u1}(\frac{\partial R_1}{\partial q_1} - \frac{\partial R_2}{\partial q_2}) = 0$$

---

[29] So, for example, we can have a situation where the unit operator receives a minority share. This is not commonly observed in unitization agreements where the firm with the largest share is usually appointed the unit operator.

Let the optimal shares be $(\alpha^{u1}, 1-\alpha^{u1})$, and the maximum *ex ante* expected surplus with M_1 as the unit operator be $\Pi^{u1}(\alpha^{u1})$. In a similar fashion, the entire problem can be repeated for M_2 as the unit operator. This leads to shares $(\alpha^{u2}, 1-\alpha^{u2})$, and an *ex ante* surplus of $\Pi^{u2}(\alpha^{u2})$. The maximum surplus attainable from a unitization agreement, and hence the choice of unit operator, solves:

(15) $\Pi^{u} = \max\{\Pi^{u1}(\alpha^{u1}), \Pi^{u2}(\alpha^{u2})\}$

Finally, we note certain features regarding the choice of the unit operator and the shares of oil that emerge in this model. Clearly the choice of the unit operator rests on the non-contractible benefits and costs from operating a reservoir, and not the contractible variables[30]. From (10) we can see that the maximization problem facing the unit operator at $t=2$ when there is agreement between the firms does not depend on which firm is the unit operator. If $R_k \neq R_l$, however, the choice of the unit operator affects the disagreement situation described in equation (8). Moreover, the choice of the unit operator is also determined by the non-contractible benefits and costs of exerting efforts (see (12a)). It is apparent that the contractible benefits V_A and V_B are not influential in determining the choice of the unit operator because both managers have the incentive to address the common pool problem and maximize the profits of operating the reservoir.

In a similar manner, we see that the choice of the optimal α depends on the fact that firms receive some non-contractible benefits from the oil extracted from the reservoir ($R_k(q_k)$)[31]. It is evident that if we assume $R_k(q_k)=0$ for all quantities of oil, any sharing rule can emerge at equilibrium. Given that these non-contractible benefits do exist[32], however, the choice of α affects the extraction decisions of the unit operator when there is a

[30] This is an entirely anticipated result. Recall from our discussion from Section 1 that the wells are assets and the selection of the unit operator determines which firm has residual control rights over the operation of the well. As the incomplete contract literature emphasizes, the allocation of residual control rights are determined by non-contractible benefits and costs.

[31] Recall that the oil is also an asset. As is indicated in the previous footnote, ownership of oil (which is dictated by the sharing rule) will be determined by non-contractible variables.

[32] Indeed, we may ask the question why real world unitization contracts stop short of marketing the oil extracted from a reservoir, that is, why they allocate shares of oil, and not shares of profits from the sale of oil. A simple answer would be that oil has value over and above its market value, which is the intuition behind R_k in this paper.

disagreement between the firms (see (8a)), the disagreement payoffs of the two firms ((9a) and (9b)), and hence the ex post bargaining strength of the firms. This, in turn, feeds back to the problem at $t = 1$, and affects the unit-operator's incentives to exert effort. The choice of α, therefore, is driven not only by the desire to maximize the non-contractible benefits of oil, but also by the necessity of providing the unit-operator with the incentives to exert effort. The result of this is evident by comparing the efficient distribution of oil (which maximizes the non-contractible benefits of oil; see (2b)), with that which would be chosen in a unitization agreement (see (14)).

3.4 Unitization versus competitive extraction

In the next section, we look at the characteristics of the various equilibrium outcomes in greater detail using specialized functional forms. To conclude this section, however, it is worthwhile to point out the main trade-offs involved in the choice between unitization and competitive extraction in this model. These are summarized below.

(i) The common pool problem: From (2a), we see that *ex post* efficiency requires that the effect that extracting oil from a given well has on the performance of the other well be taken into account. Equation (10a) indicates that unitization does take this into account, while (5c) and (5d) indicate that competitive extraction does not. The inefficiency of competitive extraction arises from this feature.

(ii) Non-contractible benefits and incomplete contracts: The contract signed in the unitization agreement is incomplete, and agents have to bargain over the division of surplus at $t = 2$. The anticipation of *ex post* bargaining distorts the incentives of the unit operator to invest at $t = 1$ (see (12a)), which leads to inefficient quantities of oil being extracted at $t = 2$. This describes the inefficiency of unitization agreements. Competitive extraction does not suffer from this inefficiency due to the absence of co-operation between firms, and hence the absence of a contract.

(iii) Substituting two managers under competitive extraction for one under unitization: Since, in reality, the non unit-operators transfer control rights to the unit operator and cease active involvement in the day-to-day operations of the oil extraction, we assumed that $i_k = 0$ for the non unit-operator. This implies that the non unit-operator fails to receive the non-contractible benefits $B_k(i_k)$-- benefits that she would have received with competitive extraction. This creates a bias against unitization which can be (artificially) inflated to establish the superiority of competitive extraction. Moreover, since any given level of effort on the two wells is exerted by one manager under unitization, while it is split between two managers under competitive extraction, sufficiently convex cost functions $(C_k(i_k))$ can deepen the bias. Implicitly, the model assumes that the unit operator cannot hire another manager, which is unrealistic. In the next section we remove this bias by assuming that $B_k(i_k) = 0$ and that $C_k(i_k)$ is linear. This focuses attention on what we believe are the main economic trade offs in the choice between competitive extraction and unitization: the common pool problem, and the distortion of efforts due to *ex post* bargaining when contracts are incomplete.

4. The failure of unitization

Section 3 outlined the solution to the model using generalized functional forms. In order to simplify the analysis and highlight the important characteristics of the trade off between unitization and competitive extraction, in this section we work with a symmetric model using specific functional forms. This is summarized in Assumption 5 below.

Assumption 5: $\forall m, n \in \{A, B\}, m \neq n$ and $\forall k, l \in \{1,2\}$, suppose the following hold.

(i) $\forall m, n \in \{A, B\}, m \neq n$:

$$V_m(q_m, q_n, i_m) = pq_m - (\beta q_m + \tfrac{1}{2}q_m^2 + \lambda q_m q_n - \sigma i_m^{1/2} q_m),$$

where $p > 0, \beta \geq 0, \sigma > 0$ and $\lambda \in (0, \tfrac{1}{2})$

(ii) $R_k(q_k) = q_k - \tfrac{1}{2}\phi q_k^2$, $\forall k \in \{1,2\}$, where $\phi \in (0, \tfrac{1}{2})$

(iii) $B_k(i_k) = 0$, $\forall k \in \{1,2\}$

(iv) $C_k(i_k) = \delta i_k$, $\forall k \in \{1,2\}$, where $\delta > 0$

In the model characterized by Assumption 5, the parameter λ captures the importance of the common pool problem in our model, while ϕ parameterizes the importance of non-contractible benefits, and the impact of contractual incompleteness; together they parameterize the two important factors determining the choice between unitization and competitive extraction (see Section 3.4).

Since $\frac{\partial^2 C_m(q_m, q_n, i_m)}{\partial q_m \partial q_n} = \lambda$, an increase in λ increases the marginal cost of extracting oil from well m when an additional unit of oil is extracted from well n, and therefore represents a worsening of the common pool problem. It is evident from the analysis in Section 3.2 that if $\lambda = 0$, competitive extraction is efficient[33]; we therefore omit this case, and focus on $\lambda > 0$[34]. Similarly $\frac{\partial^2 R_k}{\partial q_k^2} = -\phi$, and an increase in ϕ, *ceteris paribus*, increases the decline in the marginal non-contractible benefit to firm k of an additional unit of oil, making the marginal non-contractible benefit more sensitive to changes in the allocation rule, and thereby increasing the importance of non-contractible benefits in the analysis. In the extreme case where $\phi = 0$, the marginal non-contractible benefit is a constant (equal to 1), and is independent of the sharing rule. Given that the two firms are identical, any share $(\alpha, 1-\alpha)$ can be supported at equilibrium. In order to ensure an interior solution, we assume that $\phi > 0$[35].

Finally, it follows from the discussion towards the end of Section 3.4 that, given the fact that the two firms have identical non-contractible benefits and costs, the choice of the unit operator is immaterial. Without loss of generality, we assume that F_1 is the unit operator.

The subsequent analysis employs the following notation is employed:

(i) $\eta = p + 1 - \beta$

(ii) $\Delta^* = 1 + 2\lambda + \phi$

(iii) $\Delta^c = 1 + \lambda + \phi$

---

[33] Since the model in this section is symmetric and firms are identical, the allocation of investments among managers and the distribution of oil among the firms will be efficient as well.
[34] The upper bound of 1/2 on λ is placed in order to ensure the existence of en equilibrium.
[35] The upper bound on φ is placed to ensure that the results for unitization are well behaved.

(iv) $\kappa = \dfrac{\Delta^c}{\dfrac{(1+\phi)^2}{(1+\phi)^2 - \lambda^2}}$

(v) $\Delta^d = 1 + 2\lambda + 2\phi\alpha$

(vi) $\Delta^u = 1 + 2\lambda + 2\phi(\alpha^2 + (1-\alpha)^2)$

(vii) $\theta = \dfrac{1}{2}[\dfrac{1}{\Delta^u} + \dfrac{(2\alpha - 1)(1 + 2\lambda + 2\phi(2\alpha - 1))}{(\Delta^d)^2}]$

Using the analysis in Section 3, the solution for the three cases of efficiency, competitive extraction and unitization are described below; the details of the derivations are relegated to Appendix 1.

The efficient outcome: The equilibrium values of the variables are:

(E1) $q_A^* = q_B^* = q_1^* = q_2^* = \dfrac{\eta}{\Delta^* - \frac{\sigma^2}{2\delta}}$. Let $\dfrac{\eta}{\Delta^* - \frac{\sigma^2}{2\delta}} = q^*$.

(E2) $i_A^* = i_B^* = \dfrac{\sigma^2}{4\delta^2} q^{*2} = \dfrac{\eta^2\sigma^2}{4\delta^2}\left(\dfrac{1}{\Delta^* - \frac{\sigma^2}{2\delta}}\right)^2$ and $i_1^* + i_2^* = \dfrac{\eta^2\sigma^2}{2\delta^2}\left(\dfrac{1}{\Delta^* - \frac{\sigma^2}{2\delta}}\right)^2$

(E3) $\Pi^* = \eta q^*$

Competitive extraction: At the competitive equilibrium:

(C1) $q_A^c = q_B^c = q_1^c = q_2^c = \dfrac{\eta}{\Delta^c - \frac{\sigma^2}{2\delta}\left(\frac{(1+\phi)^2}{(1+\phi)^2 - \lambda^2}\right)}$. Let $\dfrac{\eta}{\Delta^c - \frac{\sigma^2}{2\delta}\left(\frac{(1+\phi)^2}{(1+\phi)^2 - \lambda^2}\right)} = q^c$.

(C2) $i_A^c = i_B^c = i_1^c = i_2^c = \dfrac{\sigma^2}{4\delta^2}\left(\dfrac{(1+\phi)^2}{(1+\phi)^2 - \lambda^2}\right)^2 q^{c^2}$

$= \dfrac{\eta^2\sigma^2}{4\delta^2}\left(\dfrac{1}{\kappa - \frac{\sigma^2}{2\delta}}\right)^2$

(C3) $\Pi^c = (1+\phi)q^{c^2} - 2\delta i^c$

Unitization:

For a given level of α:

(U1) $$i_A^u = i_B^u = \frac{\eta^2\sigma^2}{4\delta^2}\left(\frac{1}{\frac{1}{\theta}-\frac{\sigma^2}{2\delta}}\right)^2 = i^u$$

At $t = 1$, the investment by the unit operator is $i_A^u + i_B^u = 2i^u$.

(U2) When there is disagreement during the bargaining process at $t = 2$:

$$q_A^d = q_B^d = \frac{\eta + \sigma i^{u^{1/2}}}{\Delta^d} = q^d$$

$q_1^d = 2\alpha q^d$ and $q_2^d = 2(1-\alpha)q^d$

When there is agreement at $t = 2$:

$$q_A^u = q_B^u = \frac{\eta + \sigma i^{u^{1/2}}}{\Delta^u} = q^u$$

$q_1^u = 2\alpha q^u$ and $q_2^u = 2(1-\alpha)q^u$

For an interior solution, the optimal level of α solves:

(U3) $$-\phi(2\alpha-1)(q_A^u + q_B^u)^2 + (\tfrac{1}{2}\sigma q_A^u i_A^{u^{-1/2}} - \delta)\frac{\partial i_A^u}{\partial \alpha} + (\tfrac{1}{2}\sigma q_B^u i_B^{u^{-1/2}} - \delta)\frac{\partial i_B^u}{\partial \alpha} = 0$$

The total surplus under unitization is:

(U4) $$\Pi^u = (1+2\lambda)q^{u^2} - 2\delta i^u$$

We are now in a position to state the main results of the model. These are summarized by (P1) – (P5) below. The workings behind these results can be found in Appendix 1.

(P1) Competitive extraction results in over-drilling at $t = 2$. That is, $q^c > q^*$.

(P2) Competitive extraction results in over-investment at $t = 1$. That is, $i^c > i^*$.

(P3) Unitization results in under-investment at $t = 1$, for all sharing rules $(\alpha, 1-\alpha)$ that can be agreed upon at $t = 0$. In other words, $i^u < i^*, \forall \alpha \in [0,1]$.

(P4) Unitization results in under-drilling at $t = 2$. That is, $q^u < q^*$, for all sharing rules $(\alpha, 1-\alpha)$ that can be agreed upon at $t = 0$.

(P5) For an interior solution, $\alpha^u \in (\frac{1}{2}, 1)$.

(P1) captures the well-known empirical finding that in the presence of a common pool, competitive extraction leads to excessive drilling, and consequently, rent dissipation. (P2) indicates that, in our model, the *ex ante* inefficiency of competitive extraction imposed by the common pool problem is that firms may be induced to undertake excessive investments *ex ante* due to strategic effects.

While the problems with competitive extraction are well documented, unitization is generally perceived by economists to be a surplus enhancing common property solution that mitigates the ill effects of competitive extraction. Yet, the puzzle confronting the oil industry is that firms are often reluctant to enter voluntary unitization agreements. (P3) and (P4) suggest an alternative explanation to those put forward by the existing literature (see Section 1) to account for this fact. Specifically, we question the presumption that unitization is efficient and surplus enhancing relative to competitive extraction. As (P3) and (P4) indicate, unitization has inefficiencies of its own which arise from the anticipation of *ex post* bargaining: in our model, these are represented by the tendency for the unit operator to under-drill *ex post*, and to under-invest *ex ante*[36].

Given that both organizational modes involve inefficiencies, it may not always be the case that unitization is surplus enhancing relative to competitive extraction. In order to show this, we compute the surplus in both cases for specific values of the parameters. In particular, we assume that $\sigma = \delta = p = 1$ and that $\beta = 0$. This implies that $\eta = 2$. The total surpluses resulting from efficiency, competitive extraction and unitization for 25 alternative cases where the parameters λ and ϕ take values in the interval $(0, \frac{1}{2})$ are shown in Tables 3.1 to 3.3 below. The shaded cells in Table 3.2 indicate the parameter values for which the

[36] Given the specific functional forms employed, one may well question the robustness of the results. Using general functional forms, it can be shown (though we do not attempt to do so here) that when contracts are incomplete, unitization will always involve some inefficiency, though it may not be the under-investment and under-drilling that appears with our specialized model. Similarly, with a common pool of oil, competitive extraction will generally result in over-drilling, though the impact on ex ante investments is less certain.

surplus under competitive extraction exceeds that under unitization; the shaded cells in Table 3.3 indicate the reverse. Appendix 2 provides values of the other variables of interest: the optimal shares under the unitization agreement as well as quantities of oil extracted, investments and costs for the various cases.

Table 3.1: The Efficient Outcome (Π^*)

	ϕ=.05	ϕ=.1	ϕ=.2	ϕ=.3	ϕ=.4
λ=.1	5.3333	5	4.4444	4	3.6364
λ=.2	4.2105	4	3.6364	3.3333	3.0769
λ=.3	3.4783	3.3333	3.0769	2.8571	2.6667
λ=.4	2.963	2.8571	2.6667	2.5	2.3529
λ=.45	2.7586	2.6667	2.5	2.3529	2.2222

Table 3.2: Competitive extraction (Π^c)

	ϕ=.05	ϕ=.1	ϕ=.2	ϕ=.3	ϕ=.4
λ=.1	5.1929	4.8877	4.3693	3.9472	3.5979
λ=.2	3.8281	3.6886	3.4214	3.1784	2.9614
λ=.3	2.818	2.7941	2.7015	2.5839	2.4608
λ=.4	1.9552	2.0436	2.1077	2.0952	2.0481
λ=.45	1.5219	1.6807	1.8334	1.8743	1.8635

Table 3.3: Unitization (Π^u)

	ϕ=.05	ϕ=.1	ϕ=.2	ϕ=.3	ϕ=.4
λ=.1	4.6638	4.0785	3.3095	2.7748	2.3692
λ=.2	3.783	3.3891	2.8433	2.4451	2.1319
λ=.3	3.1818	2.8988	2.4916	2.1836	1.9348
λ=.4	2.7454	2.5323	2.2169	1.9718	1.7695
λ=.45	2.5692	2.3818	2.101	1.8803	1.6966

As anticipated, competitive extraction results in a higher surplus than unitization with diminishing importance of the common pool problem (lower λ), and increasing importance of non-contractible benefits (higher ϕ), which justifies our claim that unitization is not always surplus enhancing when contracts are incomplete. The tables in Appendix 2 describe the inefficiencies associated with competitive extraction and unitization in terms of quantities extracted and investments. The Appendix also shows that the optimal sharing rule under unitization is relatively sensitive to changes in ϕ for a given λ, and relatively insensitive to changes in λ for a given ϕ . This confirms our intuition that the sharing rule is dictated more by non-contractible variables than by the magnitude of the common pool problem.

Finally, we note that (P5) indicates that with symmetry, the unit operator is given a majority share in the unitization agreement. This conforms to reality where the largest share holder is often made the unit operator, but reverses the direction of the intuition by suggesting that (when firms are fairly homogenous) giving the unit operator a significant share is important in ensuring that the operator receives incentives to exert effort.

5. Extensions and conclusions

While we have focused attention on evaluating costs and benefits of unitization relative to competitive extraction, the model presented in this paper can be extended to analyze other issues related to unitization; below we provide a brief discussion on these issues.

Our model provides a framework to evaluate other organizational modes of extracting oil, such as lease consolidation. Specifically, lease consolidation can be interpreted as a special case of a unitization agreement where one firm receives all the share of oil ($\alpha = 1$ or $\alpha = 0$; that is there exists a corner solution). Our model suggests that this is likely to occur when either non-contractible benefits of oil are non-existent, or when marginal non-contractible benefits are insensitive to changes in the sharing rule. To see the latter, consider an asymmetric version of the model in Section 4 where $R_k = \omega_k q_k$. In this situation, it is evident that surplus can be maximized by giving all the oil to the firm with the higher ω_k.

McDonald (1971) outlines the role of structural advantage in the failure of unitization agreements (see Section 1), arising from the non-homogeneity of the reservoir. The notion of structural advantage can be handled by making the following alterations to the model in Section 4:

$$V_m(q_m, q_n, i_m) = pq_m - (\beta q_m + \tfrac{1}{2} q_m^2 + 2\lambda\xi q_m q_n - \sigma i_m^{1/2} q_m)$$

$$V_n(q_m, q_n, i_m) = pq_n - (\beta q_n + \tfrac{1}{2} q_n^2 + 2\lambda(1-\xi) q_m q_n - \sigma i_n^{1/2} q_n)$$

Here, the parameter ξ captures the effect of structural advantage. For example, suppose that $\xi = 0$; in such a situation, the cost of extracting oil from well *m* is independent of the amount of oil extracted from well *n*, and can be interpreted as a scenario where the oil immediately gravitates towards well *m*, forcing well *n* to bear the entire burden of the common pool problem. In short, well *m* will have a structural advantage when $\xi \in [0, \frac{1}{2})$, well *n* will have a similar advantage when $\xi \in (\frac{1}{2}, 1]$, and neither well does so when $\xi = \frac{1}{2}$, which is the case analyzed in Section 4.

Assumption 1(ii) in Section 2 reflected a feature found in reality that firms in a unitization agreement receive a share of oil equal to the share of operating costs it pays. Libecap and Smith (1999) argue that such a feature in a unitization agreement makes the contract incentive compatible and that "...the allocation formula makes each party a claimant to the unit's net profits and as such, motivates them to support a production plan that maximizes unit profits". Our analysis indicates that while this argument may be valid if there are no non-contractible benefits of oil, it breaks down when firms place a value on oil above its market value (in other words, Libecap and Smith's argument involves equating the share of oil to the share of profits, which is true only if non-contractible benefits are zero). Specifically, suppose the unit operator were to receive a share α_o of the oil and was required to bear α_c of the costs. In our model of Section 4, the *ex post* under-drilling of the reservoir, and *ex ante* under-investment that results when $\alpha_o = \alpha_c$ can be reduced by giving the unit operator a larger share of oil than costs, which would increase the unit operator's *ex post* bargaining power and increase her *ex ante* investments. Clearly, adding the extra degree of freedom due to the additional choice variable cannot decrease the surplus under unitization.

Finally, adding another stage to the model between $t = 0$ and $t = 1$ (say $t = 1/2$) permits a dynamic version of the model, where exploratory investments can be can be made at $t = 1/2$, with a distinct exploratory stage and drilling stage. This allows for a more realistic description of the oil industry, and an examination of the impact of the common pool problem and incomplete contracts on exploratory activity.

In conclusion, this paper has attempted to question the existing belief that unitization is unequivocally superior to competitive extraction for exploiting a reservoir of oil. In particular, we have shown that when contracts are incomplete, the anticipation of *ex post* bargaining may impose inefficiencies on unitization that compete with the inefficiencies associated with competitive extraction due to the common pool problem, and the outcome can be a lower surplus under unitization. In doing so, we have attempted to put forward an alternative explanation to the puzzle that oil firms are often reluctant to voluntarily enter unitization agreements.

Appendix 1

This Appendix derives the results in Section 4, based on the solutions to the general model in Section 3. In all cases, an interior solution is assumed.

The Efficient outcome:

At t = 2, from (2b), we get that $q_1 = q_2$. Given that $q_1 + q_2 = q_A + q_B$, this implies $q_k = \frac{q_A + q_B}{2}$. From (12a), we get:

(1-X) $\eta - (q_m + 2\lambda q_n - \sigma i_m^{1/2}) + 1 - \phi(\frac{q_m + q_n}{2}) = 0, \forall m \in \{A, B\}, m \neq n$

Solving the equations implied by (1-X) gives:

(2-X) $q_m^* = \frac{1}{(1-2\lambda)\Delta^*}[\eta(1-2\lambda) + \sigma i_m^{1/2}(1+\frac{\phi}{2}) - \sigma i_n^{1/2}(2\lambda + \frac{\phi}{2})], \forall m, n \in \{A, B\}, m \neq n$

Recall that we have assumed that $B_k(i_k) = 0$. Given that $i_1 + i_2 = i_A + i_B$, equation (3a) yields:

(3-X) $i_m^{*\,1/2} = \frac{\sigma}{2\delta} q_m^*, \forall m \in \{A, B\}$

Note that given the linearity of the cost functions and the absence of any private benefit to investment, the optimal allocation of effort among managers is not uniquely determined and will satisfy the condition that $i_1^* + i_2^* = i_A^* + i_B^*$.

Plugging this back in (2-X), and using the fact that $q_k = \frac{q_A + q_B}{2}$, we get:

(E1) $q_A^* = q_B^* = q_1^* = q_2^* = \frac{\eta}{\Delta^* - \frac{\sigma^2}{2\delta}} = q^*$

Inserting this in (3-X) and squaring the result gives:

(E2) $i_A^* = i_B^* = \frac{\eta^2\sigma^2}{4\delta^2}\left(\frac{1}{\Delta^* - \frac{\sigma^2}{2\delta}}\right)^2$ and $i_1^* + i_2^* = \frac{\eta^2\sigma^2}{2\delta^2}\left(\frac{1}{\Delta^* - \frac{\sigma^2}{2\delta}}\right)^2$

Competitive extraction:

With competitive extraction, by definition, $q_A = q_1, q_B = q_2, i_A = i_1$ and $i_B = i_2$. At t = 2, solving (5c) and (5d), we get:

$$\text{(4-X)}\quad q_m^c = \frac{1}{(1+\phi)^2 - \lambda^2}[\eta(1+\phi-\lambda) + \sigma(1+\phi)i_m^{1/2} - \sigma\lambda i_n^{1/2}], \forall m,n \in \{A,B\}, m \neq n$$

Solving (7c) and (7d) yields:

$$\text{(5-X)}\quad i_m^{c\,1/2} = \frac{\sigma}{2\delta}\left(\frac{(1+\phi)^2}{(1+\phi)^2 - \lambda^2}\right) q_m^c, \forall m \in \{A,B\}$$

Plugging this in (4-X), and given the fact that $q_A = q_1, q_B = q_2$, we get:

$$\text{(C1)}\quad q_A^c = q_B^c = q_1^c = q_2^c = \frac{\eta}{\Delta^c - \frac{\sigma^2}{2\delta}\left(\frac{(1+\phi)^2}{(1+\phi)^2-\lambda^2}\right)} = q^c$$

Using (C1) and (5-X) and squaring:

$$\text{(6-X)}\quad i_m^c = \frac{\sigma^2}{4\delta^2}\left(\frac{(1+\phi)^2}{(1+\phi)^2 - \lambda^2}\right)^2 \left(\frac{\eta}{\Delta^c - \frac{\sigma^2}{2\delta}\left(\frac{(1+\phi)^2}{(1+\phi)^2-\lambda^2}\right)}\right)^2, \forall m \in \{A,B\}$$

Simplifying, and using the notation that $\kappa = \dfrac{\Delta^c}{\frac{(1+\phi)^2}{(1+\phi)^2 - \lambda^2}}$ gives:

$$\text{(C2)}\quad i_A^c = i_B^c = i_1^c = i_2^c = \frac{\sigma^2}{4\delta^2}\left(\frac{1}{\kappa - \frac{\sigma^2}{2\delta}}\right)^2$$

Unitization:

Since the firms are identical, and the model is symmetric, either firm can be chosen the unit operator. In what follows, therefore, we do not distinguish a specific unit operator with a superscript as in Section 3.

If firms disagree during bargaining at t = 2, solving (8a), we get:

$$\text{(7-X)}\quad q_m^d = \frac{1}{(1-2\lambda)\Delta^d}[\eta(1-2\lambda) + \sigma(1+\phi\alpha)i_A^{1/2} - \sigma(2\lambda+\phi\alpha)i_B^{1/2}], \forall m \in \{A,B\}, m \neq n$$

On the other hand, agreement at $t = 2$ results in (10a), which in our current model results in:

$$(8\text{-X})\ q_m^u = \frac{1}{(1-2\lambda)\Delta^u}[\eta(1-2\lambda)+\sigma(1+\phi\hat{\alpha})i_A^{1/2} - \sigma(2\lambda+\phi\hat{\alpha})i_B^{1/2}], \forall m \in \{A,B\}, m \neq n$$

Here we have used the notation that $\hat{\alpha} = \alpha^2 + (1-\alpha)^2$.

At $t = 1$, the unit operator (which could be either firm) solves (12), which gives the first order condition, $\forall m \in \{A,B\}$:

$$(9\text{-X})\ \frac{1}{2}\{[\tfrac{1}{2}\sigma i_m^{-1/2}(q_m^u + (2\alpha-1)q_m^d)] - \phi(2\alpha-1)(1-\alpha)(q_m^d + q_n^d)(\frac{\partial q_m^d}{\partial i_m} + \frac{\partial q_n^d}{i_m})\} = \delta$$

Differentiating (7-X) with respect to i_m:

$$(10\text{-X})\ \frac{\partial q_m^d}{\partial i_m} + \frac{\partial q_n^d}{\partial i_m} = \frac{\sigma i_m^{-1/2}}{2\Delta^d}$$

We can exploit the symmetric nature of the model to simplify the analysis at this stage, and note that note that $i_m = i_n$ at the optimal investment levels. This implies that $q_m^d = q_n^d$, and so $q_m^d + q_n^d = 2q_m^d$. Using this and (10-X) in equation (9-X):

$$(11\text{-X})\ \tfrac{1}{2}\sigma i_m^{-1/2}(q_m^u + (2\alpha-1)q_m^d) - \frac{\phi(2\alpha-1)(1-\alpha)\sigma i_m^{-1/2} q_m^d}{\Delta^d} = 2\delta$$

$$\Rightarrow \frac{1}{2}\sigma i_m^{-1/2}[q_m^u + (2\alpha-1)q_m^d(1 - \frac{2\phi(1-\alpha)}{\Delta^d})] = 2\delta$$

$$\Rightarrow \frac{1}{2}\sigma i_m^{-1/2}[q_m^u + (2\alpha-1)q_m^d(\frac{1+2\lambda+2\phi(2\alpha-1)}{\Delta^d})] = 2\delta$$

We can use the symmetry of the problem to further simplify the expression in (11-X). Since symmetry implies that at equilibrium $i_m^u = i_n^u$, from (7-X) and (8-X), we get:

$$(12\text{-X})\ q_m^d = q_n^d = \frac{\eta + \sigma i_m^{u\,1/2}}{\Delta^d} = \frac{\eta + \sigma i_n^{u\,1/2}}{\Delta^d}, \text{ and,}$$

$$q_m^u = q_n^u = \frac{\eta + \sigma i_m^{u\,1/2}}{\Delta^u} = \frac{\eta + \sigma i_n^{u\,1/2}}{\Delta^u}$$

Using (12-X) in (11-X), we get:

$$\text{(13-X)}\quad \frac{1}{2}\sigma i_m^{u^{-1/2}}(\eta+\sigma i_m^{u^{1/2}})[\frac{1}{\Delta^u}+\frac{(2\alpha-1)(1+2\lambda+2\phi(2\alpha-1)}{(\Delta^d)^2}]=2\delta$$

Using the notation in Section 4, this reduces to:

$$\text{(14-X)}\quad \sigma i_m^{u^{-1/2}}(\eta+\sigma i_m^{u^{1/2}})\theta=2\delta$$

Simplifying this and squaring yields:

$$\text{(U1)}\qquad i_A^u=i_B^u=\frac{\eta^2\sigma^2}{4\delta^2}\left(\frac{1}{\frac{1}{\theta}-\frac{\sigma^2}{2\delta}}\right)^2=i^u$$

The result (U2) in section 4 follows from (12-X). Note that the result (U1) is true for any general α.

The total surplus under unitization, for any given α is:

$$\text{(15-X)}\quad \begin{aligned}\Pi^u &=\eta(q_A^u+q_B^u)-(\tfrac{1}{2}q_A^{u^2}+\tfrac{1}{2}q_B^{u^2}+2\lambda q_A^u q_B^u-\sigma(q_A^u i_A^{u^{1/2}}+q_B^u i_B^{u^{1/2}})\\ &-\tfrac{\phi}{2}\hat{\alpha}(q_A^u+q_B^u)^2-\delta(i_A^u+i_B^u)\end{aligned}$$

Differentiating this with respect to α, and using the envelope theorem, we get the first order condition for an interior solution:

$$\text{(U3)}\quad -\phi(2\alpha-1)(q_A^u+q_B^u)^2+(\tfrac{1}{2}\sigma q_A^u i_A^{u^{-1/2}}-\delta)\frac{\partial i_A^u}{\partial\alpha}+(\tfrac{1}{2}\sigma q_B^u i_B^{u^{-1/2}}-\delta)\frac{\partial i_B^u}{\partial\alpha}=0$$

Total surplus at t = 0

In general, the total surplus for all the three cases we have considered is:

$$\Pi=\eta(q_A+q_B)-(\tfrac{1}{2}q_A^2+\tfrac{1}{2}q_B^2+2\lambda q_A q_B-\sigma(q_A i_A^{1/2}+q_B i_B^{1/2})-\tfrac{\phi}{2}(q_1^2+q_2^2)-\delta(i_A+i_B)$$

For all three cases, $q_A=q_B=q$ and $i_A=i_B=i$ at the optimum. This implies:

$$\text{(16-X)}\quad \Pi=2\eta q-(q^2+2\lambda q^2-2\sigma q i^{1/2})-\tfrac{\phi}{2}(q_1^2+q_2^2)-2\delta i$$

For the efficient outcome, we also have that $q_1=q_2=q$. So (16-X) reduces to:

$$\text{(17-X)}\quad \Pi^*=2\eta q^*-(1+2\lambda+\phi)q^{*^2}+2\sigma q^* i^{*^{1/2}}-2\delta i^*$$

$$\Rightarrow \Pi^*=2\eta q^*-(1+2\lambda+\phi)q^{*^2}+\frac{2\sigma^2}{2\delta}q^{*^2}-\frac{\sigma^2}{2\delta}q^{*^2}\ \text{(using (3-X))}$$

Therefore:

$$\Pi^* = 2\eta q^* - (1 + 2\lambda + \phi - \frac{\sigma^2}{2\delta})q^{*2}$$

$$\Rightarrow \Pi^* = 2\eta q^* - \eta q^* \text{ (using (E1)).}$$

We thus get:

(E3) $\Pi^* = \eta q^*$

For competitive extraction we have that $q_1 = q_2 = q$ as well. This implies:

(18-X) $\Pi^c = 2\eta q^c - (1 + 2\lambda + \phi)q^{c^2} + 2\sigma q^c i^{c^{1/2}} - 2\delta i^c$

$$\Rightarrow \Pi^c = 2q^c(\eta + \sigma i^{c^{1/2}}) - (1 + 2\lambda + \phi)q^{c^2} - 2\delta i^c$$

Using the fact that $i_m^c = i_n^c$ in (4-X), we get that $q^c = \frac{\eta + \sigma i^{c^{1/2}}}{\Delta^c}$. Inserting this in (18-X):

$$\Pi^c = 2\Delta^c q^{c^2} - (1 + 2\lambda + \phi)q^{c^2} - 2\delta i^c$$

$$\Rightarrow \Pi^c = (2\Delta^c - (1 + 2\lambda + \phi))q^{c^2} - 2\delta i^c$$

Simplifying, we get:

(C3) $\Pi^c = (1 + \phi)q^{c^2} - 2\delta i^c$

For unitization, unlike the other two cases, $q_1 \neq q_2$ (unless $\alpha = 1/2$). So:

$$\Pi^u = 2\eta q^u - (1 + 2\lambda + 4\phi\hat{\alpha})q^{u^2} + 2\sigma q^u i^{u^{1/2}} - 2\delta i^u$$

$$\Rightarrow \Pi^u = 2q^u(\eta + \sigma i^{u^{1/2}}) - (1 + 2\lambda + 4\phi\hat{\alpha})q^{u^2} - 2\delta i^u$$

$$\Rightarrow \Pi^u = 2\Delta^u q^{u^2} - (1 + 2\lambda + 4\phi\hat{\alpha})q^{u^2} + 2\sigma q^u i^{u^{1/2}} - 2\delta i^u$$

$$\Rightarrow \Pi^u = (2\Delta^u - (1 + 2\lambda + 4\phi\hat{\alpha}))q^{u^2} + 2\sigma q^u i^{u^{1/2}} - 2\delta i^u$$

Simplifying, we obtain:

(U4) $\Pi^u = (1 + 2\lambda)q^{u^2} - 2\delta i^u$

Proofs for (P1) to (P5):

To begin with, note that:

$$\frac{(1+\phi)^2}{(1+\phi)^2-\lambda^2}>1 \Rightarrow \lambda+\frac{\sigma^2}{2\delta}(\frac{(1+\phi)^2}{(1+\phi)^2-\lambda^2}-1)>0$$

$$\Rightarrow (1+\phi+\lambda)+\lambda-\frac{\sigma^2}{2\delta}>(1+\phi+\lambda)-\frac{\sigma^2}{2\delta}(\frac{(1+\phi)^2}{(1+\phi)^2-\lambda^2})$$

$$\Rightarrow \Delta^*-\frac{\sigma^2}{2\delta}>\Delta^c-\frac{\sigma^2}{2\delta}(\frac{(1+\phi)^2}{(1+\phi)^2-\lambda^2})$$

Comparing (E1) and (C1), it follows that:

(P1) $q^c>q^*$

Comparing (5-X) with (3-X):

$$i^{c1/2}=\frac{\sigma}{2\delta}\left(\frac{(1+\phi)^2}{(1+\phi)^2-\lambda^2}\right)q^c>\frac{\sigma}{2\delta}\left(\frac{(1+\phi)^2}{(1+\phi)^2-\lambda^2}\right)q^*>\frac{\sigma}{2\delta}q^*=i^{*1/2}$$

Squaring, we obtain:

(P2) $i^c>i^*$

For the case of unitization, we begin by noting the following relationships, which simply follow from the definitions:

(19-X) (i) $\Delta^u>\Delta^*>\Delta^d$ for $\alpha<1/2$

(ii) $\Delta^u=\Delta^d=\Delta^*$ for $\alpha=1/2$

(iii) $\Delta^d>\Delta^u>\Delta^*$ for $\alpha>1/2$

(iv) $\Delta^d=\Delta^u>\Delta^*$ for $\alpha=1$

Note that $\Delta^u>\Delta^*$ when $\alpha\neq 1/2$, which implies that $\frac{1}{\Delta^u}<\frac{1}{\Delta^*}$ for all $\alpha\neq 1/2$. From (ii), we see that $\frac{1}{\Delta^u}=\frac{1}{\Delta^*}$ when $\alpha=1/2$.

Comparing (U1) with (E2), $i^u<i^*$ if and only if:

(20-X) $$\frac{1}{\frac{1}{\theta}-\frac{\sigma^2}{2\delta}}<\frac{1}{\Delta^*-\frac{\sigma^2}{2\delta}}\Rightarrow \theta<\frac{1}{\Delta^*}\Rightarrow \frac{1}{2}[\frac{1}{\Delta^u}+\frac{1}{\Delta^d}\frac{(2\alpha-1)(1+2\lambda+2\phi(2\alpha-1))}{\Delta^d}]<\frac{1}{\Delta^*}$$

Note, also, that $0 < \frac{1+2\lambda+2\phi(2\alpha-1)}{\Delta^d} \leq 1$ for all possible values of λ and α when $\phi \in (0,1/2)$, as is assumed in this model.

There are four relevant categories into which α can fall, and these are described in equation (19-X). We shall consider each in turn, and show that (20-X) holds true for all possible values of α.

<u>Case 1</u>: Suppose that $\alpha = 1$. Then, we see that $\frac{1+2\lambda+2\phi(2\alpha-1)}{\Delta^d} = 1$. Thus, the left hand side of (20-X) reduces to $\frac{1}{2}[\frac{1}{\Delta^u} + \frac{1}{\Delta^d}]$. From (19-X) part (iv), we can infer that $\frac{1}{\Delta^u} = \frac{1}{\Delta^d} < \frac{1}{\Delta^*}$, implying that (20-X) is satisfied in this case.

<u>Case 2</u>: Suppose that $\alpha = 1/2$. In this case, $(2\alpha - 1) = 0$, and (20-X) reduces to $\frac{1}{2\Delta^u} < \frac{1}{\Delta^*}$, which has to hold given (19-X).

<u>Case 3</u>: Suppose that $\alpha < 1/2$. Since $(2\alpha - 1) < 0$, it follows that:

$$\frac{1}{2}[\frac{1}{\Delta^u} + \frac{1}{\Delta^d}\frac{(2\alpha-1)(1+2\lambda+2\phi(2\alpha-1))}{\Delta^d}] < \frac{1}{2\Delta^u} < \frac{1}{\Delta^*}$$

<u>Case 4:</u> Suppose that $\alpha > 1/2$. For this case, from (19-X) we have that $\frac{1}{\Delta^d} < \frac{1}{\Delta^u} < \frac{1}{\Delta^*}$.

Since $\frac{1+2\lambda+2\phi(2\alpha-1)}{\Delta^d} < 1$ is also true when $\alpha > 1/2$. This implies that:

$$\frac{1}{2}[\frac{1}{\Delta^u} + \frac{1}{\Delta^d}\frac{(2\alpha-1)(1+2\lambda+2\phi(2\alpha-1))}{\Delta^d}] < \frac{1}{2}[\frac{1}{\Delta^u} + \frac{1}{\Delta^d}] < \frac{1}{\Delta^*}$$

Hence, condition (20-X) holds for all possible values of α, and so:

(P3) $i^u < i^*, \forall \alpha \in [0,1]$

Next, from (2-X), it is evident that $q^* = \frac{(\eta + \sigma i^{*1/2})}{\Delta^*}$, while from (12-X) we obtain that $q^u = \frac{\eta + \sigma i^{u1/2}}{\Delta^u}$. Result (P3) and (19-X) indicate that $\forall \alpha$, $i^u < i^*$ and $\Delta^u \geq \Delta^*$. Together, these imply:

(P4) $\quad q^u < q^*, \forall \alpha \in [0,1]$

Finally, we show that for an interior solution, it must be true that $\alpha \in (\frac{1}{2},1)$. In other words, we shall prove that the equality in (U3) can be satisfied only when $\alpha > 1/2$. In order to prove this, we must first show that $\frac{\partial i_m^u}{\partial \alpha} > 0$ for $\alpha \leq 1/2$. To see this, note that from (U1), we have:

$$i_m^u = \frac{\eta^2 \sigma^2}{4\delta^2}\left(\frac{1}{\frac{1}{\theta} - \frac{\sigma^2}{2\delta}}\right)^2 \Rightarrow sign[\frac{\partial i_m^u}{\partial \alpha}] = sign[\frac{\partial \theta}{\partial \alpha}]$$

Differentiating θ with respect to α, we get:

$$\text{(21-X)} \quad \begin{aligned} \frac{\partial \theta}{\partial \alpha} = \frac{1}{2}\{&-\frac{1}{\Delta^{u2}}\frac{\partial \Delta^u}{\partial \alpha} - \frac{2}{\Delta^{d3}}\frac{\partial \Delta^d}{\partial \alpha}(2\alpha-1)(1+2\lambda+2\phi(2\alpha-1)) \\ &+ \frac{1}{\Delta^{d2}}[2(1+2\lambda+2\phi(2\alpha-1))+(2\alpha-1)4\phi)]\} \end{aligned}$$

Differentiating Δ^u and Δ^d with respect to α:

$$\frac{\partial \Delta^u}{\partial \alpha} = 4\phi(2\alpha-1) \text{ and } \frac{\partial \Delta^d}{\partial \alpha} = 2\phi$$

Inserting this in (21-X), and rearranging:

$$\frac{\partial \theta}{\partial \alpha} = -\frac{2\phi(2\alpha-1)}{\Delta^{u2}} + \frac{(1+2\lambda+2\phi(2\alpha-1))}{\Delta^{d2}}(1-\frac{2\phi(2\alpha-1)}{\Delta^d}) + \frac{(2\alpha-1)2\phi}{\Delta^{d2}}$$

$$\Rightarrow \frac{\partial \theta}{\partial \alpha} = -\frac{2\phi(2\alpha-1)}{\Delta^{u2}} + \frac{(1+2\lambda+2\phi(2\alpha-1))}{\Delta^{d2}}(\frac{1+2\lambda+2\phi(1-\alpha)}{\Delta^d}) + \frac{(2\alpha-1)2\phi}{\Delta^{d2}}$$

For $\alpha < 1/2$ we have that $-\frac{2\phi(2\alpha-1)}{\Delta^{u^2}} > 0$, and $\frac{1+2\lambda+2\phi(1-\alpha)}{\Delta^d} > 1$. The latter implies that $\frac{(1+2\lambda+2\phi(2\alpha-1))}{\Delta^{d^2}}(\frac{1+2\lambda+2\phi(1-\alpha)}{\Delta^d}) > \left|\frac{(2\alpha-1)2\phi}{\Delta^{d^2}}\right|$. Therefore, $\frac{\partial\theta}{\partial\alpha} > 0$ for $\alpha < 1/2$ and consequently, $\frac{\partial i_m^u}{\partial\alpha} > 0$ for $\alpha < 1/2$. When $\alpha = 1/2$, $\frac{\partial\theta}{\partial\alpha} = \frac{1+2\lambda}{\Delta^{d^2}} > 0$. Hence:

(22-X) $\frac{\partial i_m^u}{\partial\alpha} > 0$ for $\alpha \le 1/2$

Next, from (11-X) we see that:

$$\delta = \frac{1}{4}\sigma i_m^{-1/2}[q_m^u + (2\alpha-1)q_m^d(\frac{1+2\lambda+2\phi(2\alpha-1)}{\Delta^d})]$$

(23-X) $\Rightarrow \frac{1}{2}\sigma q_A^u i_A^{u^{-1/2}} - \delta = \frac{1}{4}\sigma i_m^{-1/2}[q_m^u - (2\alpha-1)q_m^d(\frac{1+2\lambda+2\phi(2\alpha-1)}{\Delta^d})] > 0, \forall\alpha \le 1/2$

From (U3), we have the first order condition that must be satisfied for an interior solution:

(U3) $-\phi(2\alpha-1)(q_A^u + q_B^u)^2 + (\frac{1}{2}\sigma q_A^u i_A^{u^{-1/2}} - \delta)\frac{\partial i_A^u}{\partial\alpha} + (\frac{1}{2}\sigma q_B^u i_B^{u^{-1/2}} - \delta)\frac{\partial i_B^u}{\partial\alpha} = 0$

Equations (22-X) and (23-X) indicate that the second and third terms in the left hand side of (U3) are positive. For $\alpha < 1/2$, the first term is positive as well, while for $\alpha = 1/2$ the first term reduces to zero. Thus the equality cannot be satisfied for $\alpha \le 1/2$. We must therefore have that if an interior solution exists, it must satisfy (P5).

Appendix 2

Assuming that $\sigma = \delta = p = 1, \beta = 0$ and $\eta = 2$, this appendix provides values of quantities of oil extracted, investments and costs under the efficient outcome, competitive extraction and unitization, as well as the optimal shares in the unitization agreement for the 25 cases analyzed in Section 4.

Quantity of oil extracted:

Efficient outcome (q^*)

	ϕ=.05	ϕ=.1	ϕ=.2	ϕ=.3	ϕ=.4
λ=.1	2.6667	2.5	2.2222	2	1.8182
λ=.2	2.1053	2	1.8182	1.6667	1.5385
λ=.3	1.7391	1.6667	1.5385	1.4286	1.3333
λ=.4	1.4815	1.4286	1.3333	1.25	1.1765
λ=.45	1.3793	1.3333	1.25	1.1765	1.1111

Competitive extraction (q^c)

	ϕ=.05	ϕ=.1	ϕ=.2	ϕ=.3	ϕ=.4
λ=.1	3.0987	2.8743	2.511	2.2296	2.0051
λ=.2	2.7353	2.5546	2.2581	2.0245	1.8356
λ=.3	2.4828	2.3261	2.069	1.8659	1.7008
λ=.4	2.3118	2.1649	1.9277	1.7426	1.5929
λ=.45	2.2535	2.1064	1.8723	1.6921	1.5475

Unitization (q^u)

	ϕ=.05	ϕ=.1	ϕ=.2	ϕ=.3	ϕ=.4
λ=.1	2.2851	2.0508	1.7947	1.6264	1.4955
λ=.2	1.8596	1.703	1.5233	1.4004	1.3021
λ=.3	1.5677	1.4559	1.323	1.2293	1.1529
λ=.4	1.3551	1.2713	1.1691	1.0953	1.0342
λ=.45	1.2691	1.1954	1.1048	1.0387	0.9836

Investments:

Efficient outcome (i^*)

	ϕ=.05	ϕ=.1	ϕ=.2	ϕ=.3	ϕ=.4
λ=.1	1.7778	1.5625	1.2346	1.0000	0.8265
λ=.2	1.1080	1.0000	0.8265	0.6944	0.5917
λ=.3	0.7561	0.6944	0.5917	0.5102	0.4444
λ=.4	0.5487	0.5102	0.4444	0.3906	0.3460
λ=.45	0.4756	0.4444	0.3906	0.3460	0.3086

Competitive extraction (i^c)

	ϕ=.05	ϕ=.1	ϕ=.2	ϕ=.3	ϕ=.4
λ=.1	2.4447	2.0999	1.5984	1.2576	1.0155
λ=.2	2.0140	1.7449	1.3486	1.0750	0.8778
λ=.3	1.8272	1.5788	1.2176	0.9711	0.7944
λ=.4	1.8283	1.5561	1.1758	0.9262	0.7521
λ=.45	1.9052	1.5999	1.1867	0.9240	0.7446

Unitization (i^u)

	ϕ=.05	ϕ=.1	ϕ=.2	ϕ=.3	ϕ=.4
λ=.1	0.8010	0.4842	0.2777	0.1997	0.1572
λ=.2	0.5291	0.3356	0.2028	0.1503	0.1209
λ=.3	0.3754	0.2462	0.1545	0.1172	0.0959
λ=.4	0.2801	0.1883	0.1216	0.0939	0.0778
λ=.45	0.2455	0.1668	0.1090	0.0848	0.0707

Cost of extracting oil at t = 2:

Efficient outcome

	ϕ=.05	ϕ=.1	ϕ=.2	ϕ=.3	ϕ=.4
λ=.1	0.7112	0.6250	0.4938	0.4000	0.3306
λ=.2	0.8865	0.8000	0.6612	0.5556	0.4734
λ=.3	0.9073	0.8334	0.7101	0.6123	0.5333
λ=.4	0.8779	0.8164	0.7111	0.6250	0.5537
λ=.45	0.8561	0.7999	0.7031	0.6229	0.5555

Competitive extraction

	ϕ=.05	ϕ=.1	ϕ=.2	ϕ=.3	ϕ=.4
λ=.1	0.9162	0.7918	0.6085	0.4823	0.3917
λ=.2	1.3555	1.1937	0.9470	0.7700	0.6388
λ=.3	1.5753	1.4058	1.1416	0.9466	0.7982
λ=.4	1.6841	1.5175	1.2541	1.0559	0.9021
λ=.45	1.7139	1.5507	1.2906	1.0935	0.9397

Unitization

	ϕ=.05	ϕ=.1	ϕ=.2	ϕ=.3	ϕ=.4
λ=.1	1.0879	1.0965	0.9868	0.8603	0.7489
λ=.2	1.0681	1.0436	0.9384	0.8299	0.7341
λ=.3	1.0057	0.9733	0.8802	0.7882	0.7064
λ=.4	0.9355	0.9029	0.8224	0.7441	0.6742
λ=.45	0.9013	0.8694	0.7948	0.7225	0.6574

Optimal shares in the unitization agreement:

Share of the unit operator, Firm 1 (α^u)

	ϕ=.05	ϕ=.1	ϕ=.2	ϕ=.3	ϕ=.4
λ=.1	0.7908	0.6906	0.6004	0.5611	0.5399
λ=.2	0.7888	0.6908	0.6030	0.5644	0.5433
λ=.3	0.7873	0.6910	0.6050	0.5669	0.5460
λ=.4	0.7861	0.6911	0.6065	0.5689	0.5481
λ=.45	0.7857	0.6912	0.6071	0.5697	0.5490

IV. CONCLUSIONS

This dissertation has applied the property rights theory in two contexts: the choice of the mode of entry by a firm into a foreign country, and the choice of organizational modes adopted by oil firms while exploiting a reservoir. In doing so, the dissertation attempted to extend the property rights theory by showing that more information may not be valuable when contracts are incomplete, and by incorporating non-excludable assets within the purview of the theory.

Chapter II demonstrated that familiarity with a foreign country as well as economic similarities between two countries can play important roles in the decisions of a firm internationalizing its operations. These economic factors were summarized by isolating a familiarity effect and a similarity effect. When these effects work in opposite directions, a firm may choose to enter an unfamiliar foreign country with exports rather than FDI under the same circumstances where it would have chosen to enter with the latter had the country been a familiar one.

In order to provide an explanation for the puzzle that oil firms are often reluctant to voluntarily enter unitization agreements, Chapter III examined the choice that oil firms face between exploiting a reservoir competitively or through co-operation. The chapter highlighted the fact that both choices involve inefficiencies, contrary to the existing literature in this area which presumes that unitization is efficient. Specifically, when contracts are incomplete, the anticipation of *ex post* bargaining may impose inefficiencies on unitization that compete with the inefficiencies associated with competitive extraction due to the common pool problem. The outcome can be a lower surplus under unitization.

Extensions to the primary ideas were discussed in the concluding sections of both chapters. A common element in these discussions was the need to incorporate the analysis in a dynamic framework to capture the realities of economic interaction. This is relevant not only for the models presented here, but also for the property rights theory in general, which is essentially a static framework as it stands.

Consider the interaction between General Motors and Fisher Body, an example that crops up often in the literature to illustrate the property rights theory. The property rights theory provides valuable insights into why General Motors would want to take over Fisher Body *as soon as* the two start their relationship, but fails to explain why the two existed as separate firms for some length of time before integration actually occurred. Even a perfunctory glance at the real world would indicate that firms usually exist as separate entities before they merge. The property rights theory does succeed to a large extent where the neo-classical theory fails in demarcating the boundaries of the firm. However, the attempt to establish the factors that motivate integration must address the question of why firms decide to integrate at a specific point in time, and consequently, why they had not chosen to integrate before or after. In short, the theory needs to be extended to account for a switch in property rights over time and not just at the start of a relationship. Future research in the property rights theory has to move towards developing a dynamic perspective of integration.

References

Aghion, P. and Bolton, P. (1992). 'An Incomplete Contracts Approach to Financial Contracting'. *Review of Economic Studies*, 59: 473-494.

Aghion, P. and Tirole, J. (1994). 'The Management of Innovation'. *Quarterly Journal of Economics*, 109: 1185-1209.

Aghion, P. and Tirole, J. (1997). 'Formal and Real Authority in Organizations'. *Journal of Political Economy*, 105: 1-29.

Alchian, A. and Demsetz, H. (1972). 'Production, Information Costs, and Economic Organization'. *American Economic Review*, 62: 777-95.

American Petroleum Institute (1993). *Model Forms for Voluntary and Statutory Unit Operating Agreements.*

Arrow, K.J. (1975). 'Vertical Integration and Communication'. *Bell Journal of Economics*, 6: 173-183.

Benito, G.R.G. and Girsprud, G. (1992). 'The Expansion of Foreign Direct Investments: Discrete Rational Location Choices or a Cultural Learning Process?' *Journal of International Business Studies*, 23: 461-476.

Brainard, S.L. (1993). 'A Simple Theory of Multinational Corporations and Trade with a Trade-off between Proximity and Concentration'. *Working Paper # 4269, NBER.*

Buckley, P.J. and Casson, M. (1981). 'The Optimal Timing of a Foreign Direct Investment'. *The Economic Journal*, 91: 75-87.

Buckley, P.J. and Casson, M. (1985). *The Economic Theory of the Multinational Enterprise.* London: Macmillan.

Buckley, P.J. and Roberts, B. (1982). *European Direct Investment in the USA before World War I.* London: Macmillan.

Caves, R.E. (1996). *Multinational Enterprise and Economic Analysis.* Cambridge: Cambridge University Press.

Chiu, S. (1998). 'Non-cooperative Bargaining, Hostages and Optimal Asset Ownership'. *American Economic Review*, 88: 882-901.

Chung, J. (2001). 'Essays in International Trade'. Ph.D. dissertation, Harvard University.

Coase, R. H. (1937). 'The Nature of the Firm.' *Economica*, 4: 386-405.

Cremer, J. (1995). 'Arm's Length Relationships'. *Quarterly Journal of Economics*, 110: 275-295

Davidson W.H. (1980). 'The Location of Foreign Direct Investment Activity: Country Characteristics and Experience Effects'. *Journal of International Business Studies*, 11: 9-22.

DeMeza, D. and Lockwood, B. (1998). 'Does Asset Ownership Always Motivate Managers? Outside Options and the Property Rights Theory of the Firm'. *Quarterly Journal of Economics*, 113: 361-386.

Ekholm, K. (1998). 'Proximity Advantages, Scale Economies, and the Location of Production'. In P. Braunerhjelm and K. Ekholm (eds.), *The Geography of Multinational Firms.* Boston: Kluwer Academic.

Ethier, W.J. (1986), 'The Multinational Firm'. *Quarterly Journal of Economics*, 101:806-833.

Evans, D.S and Grossman, S. (1983). 'Integration'. In D.S. Evans (ed.), *Breaking Up Bell: Essays on Industrial Organization and Regulation.* New York: North Holland.

Fischer, S. (1977). 'Long Term Contracting, Sticky Prices, and Monetary Policy: Comment'. *Journal of Monetary Economics*, 3: 317-323.

Grossman, S. and Hart, O. (1986). 'The Costs and Benefits of Ownership: A Theory of Vertical and Lateral Integration'. *Journal of Political Economy*, 94: 691-719.

Hart, O. (1995). *Firms, Contracts and Financial Structure.* Oxford University Press.

Hart, O. and Holmstrom, B. (1987). 'The Theory of Contracts'. In T.F. Bewley (ed.), *Advances in Economic Theory.* Cambridge: Cambridge University Press, 71 -155.

Hart, O. and Moore, J. (1990). 'Property Rights and the Nature of the Firm'. *Journal of Political Economy*, 98: 1119-1158.

Hart, O. and Moore, J. (1994). 'A Theory of Debt Based on the Inalienability of Human Capital'. *Quarterly Journal of Economics*, 109: 841-879.

Hart, O. and Moore, J. (1999). 'Foundations of Incomplete Contracts.' *Review of Economic Studies*, 66: 115-138.

Helpman, E. (1984). 'A Simple Theory of International Trade with Multinational Corporations'. *Journal of Political Economy*, 92: 451-471.

Helpman, E. and Kugman, P. (1985). *Market Structure and Foreign Trade: Increasing Returns, Imperfect Competition and the International Economy.* Cambridge: MIT Press.

Hirsch, S. (1976). 'An International Trade and Investment Theory of the Firm'. *Oxford Economic Papers*, 28: 258-270.

Horstmann, I.J. and Markusen , J.R. (1996). 'Exploring New Markets: Direct Investment, Contractual Relations and the Multinational Enterprise'. *International Economic Review*, 37: 1-19

Hueth, B. and Melkonian, T. (2002). 'Asset Ownership and Incentives in Agricultural Production'. *Unpublished manuscript.*

Johanson, J. and Vahlne, J-E. (1978). 'A Model for the Decision-Making Process Affecting the Pattern and Pace of the Internationalization of the Firm.' In M. Ghertman and J. Leontiades (eds.), *European Research in International Business.* New York: North Holland.

Johanson, J. and Wiiedersheim-Paul, F. (1975). 'The Internationalization of the Firm – Four Swedish Cases'. *Journal of Management Studies*, 12: 305-322.

Klein, B., Crawford, R. and Alchian, A. (1978). 'Vertical Integration, Appropriable Rents, and the Competitive Contracting Process'. *Journal of Law and Economics*, 21: 297-326.

Laffont, J-J. and Tirole, J. (1991). 'Privatization and Incentives'. *Journal of Law, Economics and Organization*, 7: 84-105.

Luostarinen, R.(1980). *The Internationalization of the Firm.* Finland: Helsinki School of Economics.

Leuck, D. (1998). 'First Possession'. In P. Newman (ed.), *The New Palgrave Dictionary of Economics and the Law.* New York: Stockton Press.

Leuck, D. and Schenewerk, P. (1996). 'An Economic Analysis of Unitized and Non-Unitized Production'. *Proceedings of the 1996 Society of Petroleum Engineers Annual Technical Conference*, 67-76.

Libecap, G.D. and Smith, J.L. (1999). 'The Self-Enforcing Provisions of Oil and Gas Unit Operating Agreements: Theory and Evidence'. *Journal of Law, Economics and Organization*, 15: 526-548.

Libecap, G.D. and Smith, J.L. (2001). 'Regulatory Remedies to the Common Pool: the Limits of Oil Field Unitization'. *Energy Journal*, 22: 1-26.

Libecap, G.D. and Wiggins, S.N. (1985). 'The Influence of Private Contractual Failure on Regulation: the Case of Oil Field Unitization'. *Journal of Political Economy*, 93: 690-714.

Macho-Stadler, I. and Perez-Castrillo, D. (1997). *An Introduction to the Economics of Information: Incentives and Contracts.* New York: Oxford University Press.

Markusen, J.R. (1984). 'Multinationals, Multi-Plant Economies, and the Gains from Trade". *Journal of International Economics*, 16: 205-226.

Markusen, J.R. and Venables, A.J. (1998). 'Multinational Firms and the New Trade Theory'. *Journal of International Economics*, 46: 183-203.

Markusen, J.R., Venables, A.J., Konan, D. and Zhang, K. (1996) 'A Unified Treatment of Horizontal Direct Investment, Vertical Direct Investment, and the Pattern of Trade in Goods and Services'. *NBER Working Paper #5696.*

Maskin, E. (1977). 'Nash Equilibrium and Welfare Optimality'. *Working Paper, MIT.* Reproduced in *Review of Economic Studies* (1999), 66 (Special Issue: Contracts): 23-38.

Maskin, E. and Tirole, J. (1999a). 'Implementation and Renegotiation'. *Review of Economic Studies*, 66: 39-56.

Maskin, E. and Tirole, J. (1999b). 'Two Remarks on the Property Rights Literature'. *Review of Economic Studies*, 66: 139-149.

McDonald, S.L. (1971). *Petroleum Conservation in the United States.* Baltimore: The John Hopkins Press.

Nicholas, S. (1983). 'Agency Contracts, Institutional Modes, and the Transition to Foreign Direct Investment by British Manufacturing Multinationals Before 1939'. *Journal of Economic History*, 43: 675-686.

Osborne, M.J. and Rubinstein, A. (1990). *Bargaining and Markets.* San Diego: Academic Press.

Rahman, M. (2001). 'U.S. Multinational Companies: Empirical Investigation of the Shares of Sales and Exports Attributable to Affiliated and Unaffiliated Firms'. Ph.D. dissertation, Iowa State University.

Reid, S. (1984). 'Market Expansion and Firm Internalization'. In E. Kaynak (ed.), *International Marketing Management*. New York: Praeger.

Riordan, M.H. (1990). 'What is Vertical Integration?'. In M. Aoki, B. Gustafsson, and O. Williamson (eds.), *The Firm as a Nexus of Treaties*. London: Sage.

Root, F. (1994). *Entry Strategies for International Markets*. New York: Maxwell Macmillan International.

Rubinstein, A. (1982). 'Perfect Equilibrium in a Bargaining Model'. *Econometrica*, 50: 97-109.

Salanie, B. (1997). *The Economics of Contracts*. Cambridge: MIT Press.

Schmidt, K. (1991). 'The Costs and Benefits of Privatization'. Working Paper, University of Bonn .

Schmidt, K. (1996). 'Incomplete Contracts and Privatization'. *European Economic Review*, 40: 569-579.

Segal, I. (1999). 'Complexity and Renegotiation: A Foundation for Incomplete Contracts". *Review of Economic Studies*, 66: 57-82

Shapiro, C. and Willig, R.D. (1990). 'Economic Rationales for the Scope of Privatization'. In E.N. Suleiman and J. Waterbury (eds.), *The Political Economy of Public Sector Reform and Privatization*. London: Westview Press, 55-87.

Shatz, H. and Venables, A.J. (2000). 'The Geography of International Investment'. In G.L. Clark, M. Feldman and S. Gertler (eds.), *The Oxford Handbook of Economic Geography*. Oxford University Press.

Smith, J.L. (1987). 'The Common Pool, Bargaining and the Rule of Capture'. *Economic Inquiry*, 25: 631-644.

Tao, Z. and Wu, C. (1997). 'On the Organization of Cooperative Research and Development: Theory and Evidence". *International Journal of Industrial Organization*, 15: 573-596.

Tirole, J. (1999). 'Incomplete Contracts: Where Do We Stand?' *Econometrica*, 67: 741-781.

Turnbull, P.W. (1987). 'A Challenge to the Stages Theory of the Internationalization Process'. In P.J. Rosson and S.D. Reid (eds.), *Managing Export Entry and Expansion.* New York: Praeger.

UNCTAD (various issues). *World Investment Report*.

Weaver, J.L. (1986). *Unitization of Oil and Gas Fields in Texas.* The John Hopkins University Press.

Wells, L.T. (1983). *Third World Multinationals: the Rise of Foreign Investment from Developing Countries*. Cambridge: MIT Press.

Whinston, M. (2001). 'Assessing the Property Rights and Transaction-Cost Theories of Firm Scope'. *American Economic Review*, 91: 184-188.

Wiggins, S.N. and Libecap, G.D. (1985). 'Oil Field Unitization: Contractual Failure in the Presence of Imperfect Information'. *American Economic* Review, 75: 368-385.

Williamson, O. (1979). 'Transaction Cost Economics: The Governance of Contractual Relations'. *Journal of Law and Economics*, 22: 233-261.

Williamson, O. (1985). *The Economic Institutions of Capitalism.* New York: Free Press.

Yu, C.J. (1990). 'The Experience Effect and Foreign Direct Investment'. *Weltwirtschaftliches Archives*, 126: 561-80

ACKNOWLEDGEMENTS

I would like to thank Dr. Harvey Lapan for his guidance at every stage in the development of this dissertation. I will always recall, with gratitude, his kindness and patience on the numerous occasions I faltered. I would also like to thank my committee members Dr. Sam DeMarie, Dr. Brent Hueth, Dr. John Schroeter and Dr. Jinhua Zhao for helpful comments and suggestions, and Dr. Shira Lewin for introducing me to the wide world of contract theory. I suspect that this dissertation would never have been completed but for Dr. Arne Hallam's generous offer to provide financial support for my return to Iowa State University after a brief spell in Oregon. Last, but by no means least, I thank my parents, my brother Vishnu and sister-in-law Gina for their unwavering support and encouragement, and my friends for making Ames a happy town.